AD-A065 174 STANFORD UNIV CALIF DEPT OF OPERATIONS RESEARCH F/G 12/1
SOME PROBLEMS IN THE OPTIMAL CONTROL OF DIFFUSIONS.(U)
DEC 78 D SHENG N00014-76-C-0418
UNCLASSIFIED TR-84 NL

1 OF 2
ADA
065174

SOME PROBLEMS IN THE OPTIMAL CONTROL OF DIFFUSIONS

BY

DIANE DWAN SHENG

TECHNICAL REPORT NO. 84

DECEMBER 1978

PREPARED UNDER CONTRACT
N00014-76-C-0418 (NR-047-061)
FOR THE OFFICE OF NAVAL RESEARCH

Frederick S. Hillier, Project Director

This research was supported in part by National Science Foundation Grant ENG 75-14847 Department of Operations Research, Stanford University and issued as Technical Report No. 50.

DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

ACKNOWLEDGMENTS

It is with deepest appreciation that I thank Professor J. Michael Harrison, my thesis advisor. Throughout our association his thoughtful guidance and generous support have proven invaluable. I have indeed learned much from Professor Harrison and have enjoyed doing so. I am also grateful to Professor Arthur F. Veinott, Jr. for his continuous assistance and encouragement throughout my graduate studies and to Professors Chung Kai-Lai and Donald L. Iglehart for stimulating and fostering my interests in stochastic processes and applied probability.

Gail Lemmond Stein typed this manuscript and I am indebted to her efforts. Her efficiency, patience and cheerfulness have been very helpful.

This research was supported in part by the National Science Foundation and the Office of Naval Research. My thanks are extended for this financial assistance.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION AND SUMMARY

This dissertation is concerned with a class of problems in the optimal control of one-dimensional diffusion processes. We begin this chapter with an informal description of our general problem, paraphrasing the precise formulation to be given later.

We consider a controller who must employ at each point in time one of a finite number of available control modes (or actions). Her choices influence the evolution of a one-dimensional stochastic process $\{X(t);\ t \geq 0\}$ in the following way. At time zero we are given an initial non-negative level for the process X and an initial control mode. Thereafter, whenever control mode a is in use, X evolves as a Brownian Motion with drift μ_a and variance σ_a^2 that is either absorbed or instantaneously reflected at the origin. We call $X(t)$ the <u>state of the system</u> at time t. The controller is able to continuously monitor the process X but is not able to control the boundary behavior at zero.

There are costs associated with the system as follows. Whenever mode a is employed the controller continuously incurs <u>operational costs</u> at rate r_a and <u>linear holding costs</u> at rate $hX(t)$. In addition, a lump sum <u>switching cost</u> of K_{ij} is incurred instantaneously each time there is a switch from mode i to mode j. Further, in the case of an absorbing barrier there is a fixed <u>boundary cost</u> R imposed when the process X hits the origin. Finally, all costs are discounted at a positive interest rate, and the objective is to minimize expected total discounted costs over the infinite planning horizon.

We allow a very general class of control strategies, where the controller's current choice of action may depend in an arbitrary (measurable) way on past state observations and control mode selections. Intuition suggests, however, that attention can be restricted to stationary Markov policies, where changes in the control mode are dictated by only the current state and current control mode in a non-time dependent manner.

After collecting in Chapter 2 various useful preliminary results, we present in Chapter 3 a precise mathematical formulation of the control problem under study. Salient features in this development are the general definition of an admissible strategy, the precise definition of a stationary policy, the natural definition of the controlled diffusion process associated with our admissible strategies and stationary policies, and an analytical characterization of the expected total discounted costs under a stationary policy. In Chapter 4 we provide necessary and sufficient conditions for a given stationary policy to be optimal. We conjecture that there always exists a stationary policy that is optimal, but offer no general proof at this time.

The main results of this dissertation concern the control problem described above in the case of two available control modes. Optimal stationary policies and their associated expected costs are explicitly computed for such problems in Chapter 5. When the switching costs are all zero and with either absorbtion or reflection, the optimal policy dictates choice of action as a function of the current state of the system only and is characterized by a single critical number. One control mode is to be used whenever the process X is above this critical number z $(0 \leq z < \infty)$,

and the other control mode is used if X is below level z. The single critical number is shown to be the unique solution to a complicated transcendental equation.

When there is a positive switching cost for a change in the control mode, and with either boundary behavior, there again is a stationary policy that is optimal. But in this situation, the optimal policy selects actions according to a function of both current state and current control mode. This policy is characterized by two critical numbers z and Z $0 \leq z < Z \leq \infty$. One control mode is called for whenever the state of the system is above Z, and the other mode is used when the state is below z. If the process X falls between the critical numbers, the controller simply maintains the control mode currently in use. As in the case of zero switching costs, we are able to derive formulas for the calculation of these critical numbers.

Our study is related to earlier work by Mandl (1968) and Pliska (1973) on the optimal control of diffusion processes. The added feature of our formulation is the inclusion of lump sum switching costs, whereas Mandl and Pliska assume that costs are continuously incurred at a rate dependent on both current state and current action. In the following regards, however, our formulation is more specialized than those of Mandl and Pliska.

(1) There are available only a finite number of control modes.

(2) The infinitesimal mean and infinitesimal variance of our controlled process X depend on the action currently in use but not on the current state of the system.

(3) Our cost rate function has the special form $g(x,a) = hx + r_a$, when in state x and employing mode a.

In formulating and analyzing our stochastic control problem, we use the Itô approach to diffusion theory, whereas Mandl and Pliska rely heavily on the analytical theory of diffusions (and general Markov processes) associated with Feller (1954, 1957) and Dynkin (1965). Given the essential restriction (2), our approach yields a natural definition of admissible (generally non-stationary) strategies, whereas Mandl and Pliska must confine attention to stationary policies throughout their formulations. In addition, we avoid most of the complexity inherent in the Feller-Dynkin approach to diffusions by throwing all analytical issues onto Itô's lemma, a relatively simple sample path relationship.

One large area of application for our control problem with reflection is in the control of queueing systems. Consider a service facility with infinite waiting room and a controller who may choose at each point in time one of two different available servers for duty in the system. There are no specific assumptions about the form of the interarrival time distribution or the service time distributions, except that the arrival process is stationary over time and that the server on duty can not remain idle if the queue length is positive. At any time the controller's choice of server is dependent only on the current queue length and identity of the server currently on duty. For a given stationary control policy as defined in Chapter 3 here, it is shown in Rath (1975) that as a sequence of these controlled queues converges to heavy traffic conditions, a normalized sequence of the queue length process converges weakly to the

controlled reflected diffusion process formalized here. (Rath's meaning of heavy traffic conditions is that the mean service rate for each available server is approximately equal to the mean arrival rate.) He shows also that the accumulated costs without discounting arising from operating, holding, and switching costs likewise converge weakly to the respective total undiscounted costs generated by the controlled reflected diffusion process. These results can be extended to a cost structure that includes discounting, an arbitrary finite number of available servers, multiple arrival channels and multiple service slots, and thus our control problem is appropriate in studying the optimal control of a variety of queueing systems in heavy traffic.

Diffusions models, in general, and modified versions of our control problem, in particular, have also received attention in application to water reservoirs, stochastic cash management, collective risk theory, inventories and other input-output systems. Here we mention two such examples. Others are included in the bibliography. Puterman (1975) uses diffusion processes to model continuous time storage systems where in the language of our control problem there are two available control modes such that $\mu_1 > 0$ and $\mu_2 < 0$, linear holding costs, switching costs, no operational costs, and no barrier at zero (i.e., backlogging is allowed). There is also no discounting of costs over time, and under the objective of minimizing long-run average costs he investigates optimizing over the class of our two critical numbers policies where $z < 0 < Z$. Zuckerman (1977) studies a finite capacity water reservoir via our control problem where there is both an upper and lower reflecting barrier and available two control modes

such that $\mu_1 > 0$, $\mu_2 = \mu_1 - M < \mu_1$ and $\sigma_1^2 = \sigma_2^2$. (The input of water into the reservoir is a Brownian Motion process with positive drift μ_1, while the controller chooses water to be continuously released at rate 0 or M.) There are no holding costs in the Zuckerman model, and the operational and switching costs are such that $r_1 = 0$, $r_2 < 0$, $K_{21} = 0$ and $K_{12} \geq 0$. His objective is to minimize discounted costs and he allows only two critical numbers policies of the form $z = 0$.

Finally, we note here the accounting conventions used in this dissertation. Theorems, propositions, equations and definition statements are numbered consecutively within chapters, with the numbering starting anew in each chapter. Within the same chapter, theorems, propositions, equations, and definition statements are referenced simply by their number, whereas cross chapter references also require a chapter identification.

CHAPTER 2

PRELIMINARIES

In this chapter we present some terminology and preliminary results from the theory of stochastic integration. This material will be used later in defining and discussing the stochastic processes that arise in conjunction with our control problem. The reader is referred to McKean (1969), Gihman and Skorohod (1972), and Ash and Gardner (1975) for background information on stochastic integrals and stochastic differential equations.

2.1. A Variation of Itô's Lemma

We begin by introducing some notation and terminology that will be continued throughout this paper. Their relevance will be seen in Chapter 3 in the construction of a class of controlled diffusion processes and associated value functions.

Start with a probability space $(\Omega, \mathfrak{F}, P)$ on which is defined a standard (zero drift and unit variance) Brownian Motion $B = \{B(t);\ t \geq 0\}$ with $B(0) = 0$. Let $\{\mathfrak{F}_t;\ t \geq 0\}$ be the increasing family of sub-σ-fields $\mathfrak{F}_t = \mathfrak{F}\{B(u);\ 0 \leq u \leq t\}$ generated by the process B. Let H denote the set of functions $f : \Omega \times [0,\infty) \to \mathbb{R}$ such that

(1) $f(\omega,t)$ is jointly measurable in ω and t,

(2) $f(\cdot, t)$ is $\mathfrak{F}_t$ measurable for each $t \geq 0$, and

(3) $$P\{\omega : \int_0^t f^2(\omega, u)\, du < \infty\} = 1 \qquad \text{for each } t \geq 0 .$$

Elements of H are called <u>non-anticipating Brownian functions</u>, or just <u>non-anticipating functions</u> for short. We can suppress the ω notation in the above, refer to $\{f(t);\ t \geq 0\}$ as a <u>non-anticipating process</u>, and re-state (3) as

(3) $$\int_0^t f^2(u)\, du < \infty \quad \text{almost surely for each } t \geq 0 .$$

Now recall that for all $f \in H$, the stochastic integral $\int_0^t f(u)\, dB(u)$ of f with respect to the Brownian Motion B is well defined for each $t \geq 0$. The stochastic process I_f defined by

$$I_f(t) = \int_0^t f(u)\, dB(u) , \qquad t \geq 0 ,$$

is almost surely (a.s.) continuous in t. (Hereafter, we shall simply say that I_f is a continuous process.) Furthermore, if

$$\int_0^t E[f^2(u)]\, du < \infty , \qquad t \geq 0 ,$$

then we have

(4) $$E[(I_f(t))^2] = \int_0^t E[f^2(u)]du \quad \text{and} \quad E[I_f(t)] = 0, \quad t \geq 0 .$$

We shall call a process $\{X(t);\ t \geq 0\}$ an <u>Itô process</u> if it is of the form

$$(5) \qquad X(t) = X(0) + \int_0^t \mu(u)\, du + \int_0^t \sigma(u)\, dB(u)\ , \qquad t \geq 0\ ,$$

where $X(0)$ is a random variable measurable with respect to $\mathcal{F}_0$, and μ and σ belong to H. The first integral on the right hand side of (5), called the <u>drift component</u> of process X, is defined for almost all $\omega \in \Omega$ as a Lebesgue integral, so it is continuous as a function of t. The stochastic integral in (5) is called the <u>diffusion component</u> of X. Thus, X is a continuous process and $X(t)$ is $\mathcal{F}_t$ measurable for each $t \geq 0$.

The following is a statement of Itô's lemma, which is proved on pages 24-25 of Gihman and Skorohod (1972).

<u>Theorem 1</u>. Let X be an Itô process and $f : [0,\infty) \times \mathbb{R} \to \mathbb{R}$, a continuous function with continuous partial derivatives $f_t(t,x)$, $f_x(t,x)$ and $f_{xx}(t,x)$. Then the process $\{f(t, X(t));\ t \geq 0\}$ is an Itô process satisfying

$$(6) \qquad f(t, X(t)) = f(0, X(0)) + \int_0^t [f_u(u, X(u)) + \tfrac{1}{2} f_{xx}(u, X(u))\, \sigma^2(u) + f_x(u, X(u))\, \mu(u)]\, du + \int_0^t f_x(u, X(u))\, \sigma(u)\, dB(u)\ , \qquad t \geq 0\ .$$

We would like to extend Itô's lemma to a broader class of functions f and a broader class of processes X. For S an interval in $\mathbb{R}$, $C^2(S)$ is the conventional notation for the set of twice continuously differentiable functions on S. Let $C_*^2(S)$ denote the set of continuously differentiable functions $f : S \to \mathbb{R}$ such that f'' exists and is continuous at all but a finite number of points in each finite interval of S, and such that the left and right second derivatives $f''(x-)$ and $f''(x+)$ exist and are finite everywhere in S. The next theorem is the extension of Itô's lemma that we will need later.

Theorem 2. Let X and Y be a pair of processes satisfying

$$X(t) = X(0) + \int_0^t \sigma(u)\, dB(u) + Y(t) \tag{7}$$

for each $t \geq 0$, where $X(0)$ is $\mathcal{F}_0$ measurable, σ is bounded and non-anticipating, and Y is a continuous non-anticipating process of bounded variation. Let $f:[0,\infty] \times \mathbb{R} \to R$ be such that

(8) $\quad f(t,\cdot) \in C_*^2(\mathbb{R})$ for each $t \geq 0$, and

(9) $\quad f(\cdot,x)$ is continuously differentiable on $[0,\infty)$ for each $x \in \mathbb{R}$.

Then the process $\{f(t, X(t));\ t \geq 0\}$ satisfies

(8) $$f(t, X(t)) = f(0, X(0)) + \int_0^t [f_u(u, X(u)) + \frac{1}{2} f_{xx}(u, X(u)) \sigma^2(u)]du$$

$$+ \int_0^t f_x(u, X(u)) \sigma(u) \, dB(u)$$

$$+ \int_0^t f_x(u, X(u)) \, dY(u) , \qquad \text{for each} \quad t \geq 0 .$$

Remark. In familiar fashion, the process Y may simply have the form

(11) $$Y(t) = \int_0^t \mu(u) \, du , \qquad t \geq 0 ,$$

where μ is bounded and non-anticipating. In this case one clearly substitutes $\mu(u)du$ for $dY(u)$ in (10). Our principal interest, however, will be in processes Y which are continuous and of bounded variation but not absolutely continuous. (In general, our processes Y will be the sum of an absolutely continuous part of the form (11) and a continuous part of bounded variation.) Such processes arise in the representation of diffusions with reflecting barriers. Note that X defined by (7) above is not then an Itô process according to our definition (5).

Proof. The proof here relies heavily on a similar result, Theorem 2.2, found in Kunita and Watanabe (1967). Their statement of (10) is for functions $f : \mathbb{R}^N \to \mathbb{R}$ with continuous first and second order partial derivatives, as applied to $\{\mathfrak{F}_t; t \geq 0\}$ measurable processes $\{Z(t); t \geq 0\}$ of the form $Z(t) = M(t) + R(t)$, where $\{M(t); t \geq 0\}$ is a N-dimensional martingale and $\{R(t); t \geq 0\}$ is a continuous N-dimensional rectifiable process.

Here the process X-Y is a $\{\mathfrak{F}_t;\ t \geq 0\}$ measurable martingale, where additionally

$$E\left[[X(t) - Y(t) - X(\tau) + Y(\tau)]^2 | \mathfrak{F}_\tau\right] = E\left[\int_\tau^t \sigma^2(u)\, du\right].$$

Now suppose that instead of (8), we have that

$$(8') \qquad f(t,\cdot) \in C^2(\mathbb{R})$$

for each $t \geq 0$. Then we can apply the above Kunita and Watanabe theorem to processes X and Y, resulting in (10) where the last integral is defined for almost all $\omega \in \Omega$ as a Lebesgue integral.

Fix $t \geq 0$. Since X is continuous, we have that $f_u(u,x)$ and $f_x(u,x)$ are continuous and bounded on $[0,t] \times S$, where S is the finite interval $[\min_{0 \leq u \leq t} X(u), \max_{0 \leq u \leq t} X(u)]$. The second partial derivative $f_{xx}(u,x)$ is continuous on $[0,t] \times S$ except at a finite number of points in S, while the left and right second partials $f_{xx}(u, x-)$ and $f_{xx}(u, x+)$ are bounded on $[0,t] \times S$. We now construct a sequence of functions $f^n : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ such that (8') and (9) hold for each f^n, $f^n(t,x)$ converges to $f(t,x)$ pointwise on $[0,t] \times S$, and such that $f^n_u(u,x)$, $f^n_x(u,x)$ and $f^n_{xx}(u,x)$ converge almost everywhere to $f_u(u,x)$, $f_x(u,x)$ and $f_{xx}(u,x)$, respectively on $[0,t] \times S$. Therefore, we have

$$(10') \qquad f^n(t, X(t)) = f^n(0, X(0)) + \int_0^t [f_u^n(u, X(u)) + \frac{1}{2} f_{xx}^n(u,X(u))\sigma^2(u)]du$$

$$+ \int_0^t f_x^n(u, X(u))\, \sigma(u)\, dB(u) + \int_0^t f_x^n(u, X(u))\, dY(u)$$

for each n.

Using the Lebesgue Dominated Convergence Theorem we have that $f_u^n(u, X(u))$, $f_x^n(u, X(u))$ and $f_{xx}^n(u, X(u))$ converge to $f_u(u, X(u))$, $f_x(u, X(u))$ and $f_{xx}(u, X(u))$ in L^2 convergence, for each $u \in [0,t]$. Relying on the boundedness of μ and σ and once again applying the Lebesgue Dominated Convergence Theorem, we see that the Lebesgue integrals

$$\int_0^t E[[f_u^n(u, X(u)) - f_u(u, X(u))]^2]du \; ,$$

$$\int_0^t E[[f_{xx}^n(u, X(u)) - f_{xx}(u, X(u))]^2 \sigma^4(u)]du \; ,$$

$$\int_0^t E[[f_x^n(u, X(u)) - f_x(u, X(u))]^2 \sigma^2(u)]du \; ,$$

and

$$\int_0^t E[[f_x^n(u, X(u)) - f_x(u, X(u))]^2]du$$

each converge to zero as $n \to \infty$. This implies L^2 convergence to zero of the random variables

$$\int_0^t [f_u^n(u, X(u)) - f_u(u, X(u))]du ,$$

$$\int_0^t [f_{xx}^n(u, X(u)) - f_{xx}(u, X(u))] \sigma^2(u)du ,$$

$$\int_0^t [f_x^n(u, X(u)) - f_x(u, X(u))] \sigma(u) \, dB(u) ,$$

and

$$\int_0^t [f_x^n(u, X(u)) - f_x(u, X(u))] \, dY(u) ,$$

which finally gives convergence in probability of the right hand side of (10') to

$$f(0, X(0)) + \int_0^t [f_u(u, X(u)) + \tfrac{1}{2} f_{xx}(u, X(u)) \sigma^2(u))]du$$
$$+ \int_0^t f_x(u, X(u)) \sigma(u) \, dB(u) + \int_0^t f_x(u, X(u)) \, dY(u) ,$$

as desired. □

2.2. A Simple Stochastic Differential Equation

Later we will be describing controlled diffusion processes as the solutions to stochastic integral equations of the form

$$X(t) = X(0) + \int_0^t \mu(X(u))du + \int_0^t \sigma(X(u)) \, dB(u) , \quad \text{for all } t \geq 0 ,$$

where $\mu(\cdot)$ and $\sigma(\cdot)$ are piece-wise constant. Theorem 3 will be helpful in discussion of the existence and uniqueness of such solutions.

Theorem 3. Let θ, γ, α, β and s be constants such that $\gamma > 0$ and $\beta > 0$. Given $X(0)$, a $\mathfrak{F}_0$ measurable random variable, there exists a unique (in distribution) non-anticipating process X satisfying

$$(12) \qquad X(t) = X(0) + \int_0^t [\theta \chi\{X(u) > s\} + \alpha \chi\{X(u) \le s\}]du$$

$$+ \int_0^t [\gamma \chi\{X(u) > s\} + \beta \chi\{X(u) \le s\}]\, dB(u) ,$$

$$\text{for all } t \ge 0 ,$$

where $\chi\{\mathcal{E}\}$ is the indicator function of event $\mathcal{E}$.

Remark. If X is non-anticipating, then the integrands on the right hand sie of (12) are non-anticipating Brownian functions. Thus the Itô integral in (12) is well-defined and X is an Itô process.

Proof. Letting μ and σ be the following piece-wise continuous functions on $\mathbb{R}$

$$\mu(x) = \begin{cases} \alpha , & \text{if } x \le s \\ \theta , & \text{if } x > s \end{cases}$$

and

$$\sigma(x) = \begin{cases} \beta , & \text{if } x \leq s \\ \gamma , & \text{if } x > s , \end{cases}$$

we begin by applying a standard random time substitution and state transformation to process B. The time substitution is as discussed on pages 111 - 113 in Gihman and Skorohod (1972). That is, for all $t \geq 0$ define $T(t)$ by

$$t = \int_0^{T(t)} \frac{1}{\delta^2(B(u))} \, du , \tag{13}$$

where $\delta(\cdot)$ is a positive piece-wise continuous function (to be specified in a moment) such that

$$\int_0^{\infty} \frac{1}{\delta^2(B(u))} \, du = +\infty , \qquad \text{a.s.} \tag{14}$$

and

$$\int_0^{\tau} E\left[\frac{1}{\delta^2(B(u))}\right] du < \infty , \qquad \text{for all } \tau \geq 0 . \tag{15}$$

Note that $T(\cdot)$ defined by (13) is a strictly increasing and continuous one-one mapping of $[0,\infty)$ into $[0,\infty)$ such that $T(0) = 0$. Therefore, the process $\{W(t);\ t \geq 0\}$ defined by $W(t) = B(T(t))$ satisfies

$$W(t) = W(0) + \int_0^t \delta(W(u)) \, dB^*(u) , \qquad \text{for all } t \geq 0 , \tag{16}$$

where $\{B^*(t);\ t \geq 0\}$ is the Brownian Motion defined on $(\Omega, \mathfrak{F}, P)$ by

$$(17) \qquad B^*(t) = \int_0^{T(t)} \frac{1}{\delta(B(u))}\, dB(u)\ , \qquad \text{for all } t \geq 0\ .$$

(Since $T(t)$ is trivially $\mathfrak{F}_t$ measurable for all $t \geq 0$, $B^*(t)$ is also, and hence $W(t)$ is likewise properly measurable.) We now scale the state space of process W by introducing the following function on $\mathbb{R}$,

$$g(x) = \begin{cases} \displaystyle\int_0^x \exp\left\{- \int_0^y \frac{2\mu(z)}{\sigma^2(z)}\, dz\right\} dy\ , & \text{if } x \geq 0 \\[2ex] \displaystyle -\int_0^{-x} \exp\left\{\int_0^y \frac{2\mu(-z)}{\sigma^2(-z)}\, dz\right\} dy\ , & \text{if } x < 0\ . \end{cases}$$

Since g is strictly increasing and continuous in x its inverse function $f = g^{-1}$ is well-defined, and therefore the process $\{X^*(t);\ t \geq 0\}$ defined by $X^*(t) = X(0) + f(W(t))$ is $\{\mathfrak{F}_t;\ t \geq 0\}$ measurable. Itô's lemma now gives us that

$$(18) \qquad X^*(t) = X^*(0) + \int_0^t \frac{1}{2} f''(W(u))\, \delta^2(W(u))du + \int_0^t f'(W(u))\, \delta(W(u))\, dB^*(u)\ ,$$

$$\text{for all } t \geq 0\ ,$$

so if we specify δ such that $\delta^2(W(t)) = [g'(X^*(t))]^2\, \sigma^2(X^*(t))$ we will have that X^* satisfies

$$(19) \qquad X^*(t) = X^*(0) + \int_0^t \mu(X^*(u))du + \int_0^t \sigma(X^*(u))\, dB^*(u), \quad \text{for all } t \geq 0.$$

Moreover, the boundedness of functions μ and σ guarantee that X^* is non-anticipating and that $\sup\limits_{0 \leq t \leq \tau} E[[X^*(t)]^2] < \infty$ for all $\tau \geq 0$.

Thus we have shown the existence of a non-anticipating process X satisfying (12), since if there exists a non-anticipating X^* satisfying (19) for a particular Brownian Motion B^* there then exists such an X for any Brownian Motion B. This can be seen by looking at $F : C[0,\infty) \to C[0,\infty)$ a path-to-path mapping such that $X^* = F(B^*)$, and letting $X = F(B)$. ($C[0,\infty)$ denotes the set of all continuous functions on $[0,\infty)$.)

We now prove the uniqueness (in distribution) of X by showing any solution to (12) to be a diffusion process with drift coefficient $\mu(x)$ and diffusion coefficient $\sigma^2(x)$. That is, following the definition of diffusion in Breiman (1968), we need verify that X is a Feller process such that for all $\epsilon > 0$ and all $x \in \mathbb{R}$

i) $$\lim_{\Delta \downarrow 0} \int_{|y-x|>\epsilon} P(t,x,t+\Delta,dy) = 0,$$

ii) $$\lim_{\Delta \downarrow 0} \frac{1}{\Delta} \int_{|y-x|\leq\epsilon} (y-x)\, P(t,x,t+\Delta,dy) = \mu(x),$$

and

iii) $$\lim_{\Delta \downarrow 0} \frac{1}{\Delta} \int_{|y-x|\leq\epsilon} (y-x)^2\, P(t,x,t+\Delta,dy) = \sigma^2(x),$$

where $P(t,x,\tau,U)$ is the transition probability function for X ($0 \leq t \leq \tau$ and U a Borel set). Theorem 1 on page 67 of Gihman and Skorohod (1972) demonstrates immediately that X is a Feller process, and the arguments that follow on pages 68-69 will also carry over here to prove i), ii) and iii) above. □

Note that if, in the statement of the theorem, we had allowed $\gamma = \beta = 0$, no stochastic integral would have been involved. Equation (12) could then be viewed as an ordinary differential equation to be solved for each ω, but there is no guarantee that a solution exists. See page 78 of McKean (1969).

We make note also of our belief that the process X as a solution to (12) is indeed unique up to a stochastic equivalence, although we did not prove so here. Uniqueness in distribution is all that we will need later.

CHAPTER 3

THE GENERAL FORMULATION

In this chapter we formulate precisely a class of problems involving the optimal control of one-dimensional Brownian Motion. For each admissible control strategy, there is a corresponding controlled diffusion process which generates costs according to a specified mechanism. The objective is to construct an admissible control strategy which minimizes expected discounted costs over an infinite planning horizon.

The first section discusses the states of the system, the available actions, and the cost structure. Section 3.2 defines an admissible strategy and the associated controlled process. In Section 3.3 we treat the expected discounted costs corresponding to an admissible strategy and define optimality. The fourth section develops stationary policies and their expected discounted costs.

3.1. The Data of the Problem

We assume given a finite set $A = \{1, 2, \ldots, N\}$ of potential actions, or control modes, and real constants r_a, μ_a and σ_a for each $a \in A$. These constants are called the operational cost rate, drift parameter and variance parameter, respectively for mode a, and we assume that $\sigma_a > 0$ for all $a \in A$.

Also given are non-negative switching costs K_{ij} for i and j in A, which satisfy

(1) $$K_{ii} = 0 , \qquad \text{for all } i \in A ,$$

and

(2) $$K_{ij} \leq K_{ia} + K_{aj} , \qquad \text{for all } i,j,a \in A .$$

Next, there is a holding cost rate h, a boundary cost R and an interest rate $\alpha > 0$. (Note that the term "cost" is rather artificial since h, R and each of the r_a are unrestricted in sign.) Finally, we are given a boundary parameter $\lambda \in \{0,1\}$.

By way of interpretation, we imagine a controller who continuously monitors the state of a system $\{X(t);\ t \geq 0\}$ and who must employ at each point in time $t \geq 0$ some control mode $a \in A$. The state space for the problem is $S = [0,\infty)$ and the action space is the set A. When mode a is in use, the state of the system changes like a one-dimensional Brownian Motion with infinitesimal drift μ_a, infinitesimal variance σ_a^2, and either absorbtion or instantaneous reflection at the origin. If $\lambda = 0$, then the barrier at zero is absorbing, and if $\lambda = 1$, there is instantaneous reflection at the barrier.

There are costs associated with the trajectory of X that are dependent on the controls employed and the state of the system. First, costs are continuously incurred at a rate $hx + r_a$ whenever the state of the system is $x > 0$ and control mode a is in use, and at a rate $(1-\lambda)\,\alpha R$ when the state of the system is zero. Additionally, the lump sum cost K_{ij} is incurred instantaneously whenever there is a change from control mode i to mode j. And finally, all costs are discounted at interest rate $\alpha > 0$. (Thus, a cost C incurred at time t is

equivalent to a cost of $Ce^{-\alpha t}$ incurred at time zero.) These interpretations for the data of the problem will be justified by the formal definitions of the next two sections.

Based on the switching costs we can divide up the action space into disjoint equivalence classes. For each action $a \in A$, let $\theta(a) = \{j \in A : K_{aj} = K_{ja} = 0\}$. Thus $\theta(a)$ is the set of all control modes that can be reached from a and from which a can be reached without switching costs. From (2) it follows that for any a and i in A, either the sets $\theta(a)$ and $\theta(i)$ are identical or have an empty intersection. Therefore, there exist M disjoint equivalence classes, denoted $A_1, A_2, \ldots, A_M$, where $\bigcup_{k=1}^{M} A_k = A$ and $A_k = \bigcup_{a \in A_k} \theta(a) = \bigcap_{a \in A_k} \theta(a)$ for each $k \in \{1, 2, \ldots, M\}$. It costs nothing to switch in either direction among actions in any one equivalence class, and there is a positive cost to switching in some direction among actions of different equivalence classes. Furthermore, there exist non-negative costs $C_{k\ell}$ for k and ℓ in $\{1, 2, \ldots, M\}$ representing the costs of switching between equivalence classes that satisfy

(3) $\quad C_{k\ell} = K_{ij}$, for all $i \in A_k$ and all $j \in A_\ell$, for all $k, \ell \in \{1,2,\ldots,M\}$,

(4) $\quad C_{kk} = 0 \quad$ for all $k \in \{1, 2, \ldots, M\}$,

(5) $\quad C_{k\ell} \leq C_{km} + C_{m\ell} \quad$ for all $k, \ell, m \in \{1, 2, \ldots, M\}$.

We end this section with one final note regarding the generality of our switching costs. Originally, we could have allowed negative

switching costs K_{ij} if, in addition to (1) and (2), we had required that

$$K_{a(1)a(2)} + K_{a(2)a(3)} + \cdots + K_{a(p-1)a(p)} \geq 0$$

for every finite sequence $a(1), a(2), \ldots, a(p)$ selected from A where $a(1) = a(p)$. The resultant equivalence class costs then would likewise no longer be necessarily non-negative but would satisfy the above non-negative curcuit condition in addition to (3), (4) and (5). However, for the sake of simplicity, we have required that $K_{ij} \geq 0$ for all actions i and j.

3.2. Admissible Strategies

As in Section 2.1, we start with a given probability space $(\Omega, \mathfrak{F}, P)$, on which is defined a standard (zero drift and unit variance) Brownian Motion $B = \{B(t); t \geq 0\}$ with $B(0) = 0$. Let $\{\mathfrak{F}_t; t \geq 0\}$ be the increasing (and continuous) family of sub-σ-fields generated by B. We now describe what is allowed as a control strategy.

Definition. An admissible strategy is a function $\pi : \Omega \times (0,\infty) \to A$ such that

(6) $\pi(\omega,t)$ is jointly measurable in ω and t ,

(7) $\pi(\cdot,t)$ is $\mathfrak{F}_t$ measurable for all $t > 0$, and

(8) the function $\theta^* : \Omega \times (0,\infty) \rightarrow \{1, 2, 3, \ldots, M\}$ defined by $\theta^*(\omega,t) = \theta(\pi(\omega,t))$ has only finitely many discontinuities in each finite interval of time.

Hereafter we will suppress the dependence of π on ω, and so $\{\pi(t); t > 0\}$ is the process representing the action used under strategy π at each point in time $t > 0$.

We now associate with each admissible strategy π and initial state X a corresponding controlled Brownian Motion. It will of course be necessary to distinguish between problems with absorbtion $(\lambda = 0)$ and those with reflection $(\lambda = 1)$.

Theorem 1. Let π be an admissible strategy and assume $x \in S$. There exist a unique pair of non-anticipating processes X and Y which jointly satisfy

$$(9) \qquad X(t) = x + \int_0^t \mu_{\pi(u)}\, du + \int_0^t \sigma_{\pi(u)}\, dB(u) + Y(t)\ , \qquad \text{for all } t \geq 0\ ,$$

(10) $Y(\cdot)$ is continuous, non-decreasing with $Y(0) = 0$, and grows only when $X(t) = 0$.

Remark. From (9) and (10) it follows that X is continuous with $X(0) = x$. Explicit formulas for X and Y are given in the proof that follows.

Proof. By the definition of admissible strategy and the boundedness of μ_π and σ_π, μ_π and σ_π are easily non-anticipating. Hence the process $\{Z(t);\ t \geq 0\}$ uniquely defined by

$$Z(t) = x + \int_0^t \mu_{\pi(u)}\, du + \int_0^t \sigma_\pi(u)\, dB(u)\ , \qquad t \geq 0\ ,$$

is an Itô process. Since Z is continuous we have that

$$\int_0^t Z^2(u)\, du < \infty\ , \qquad \text{almost surely for each } t \geq 0\ ,$$

and so, Z is a non-anticipating process.

Let $\{Y(t);\ t \geq 0\}$ be defined by

$$(11) \qquad Y(t) = \left[-\inf_{0 \leq u \leq t} \{Z(u)\}\right]^+\ , \qquad t \geq 0\ .$$

Clearly Y is continuous, non-decreasing and non-anticipating, and $Y(0) = 0$. Now let $\{X(t);\ t \geq 0\}$ be such that

$$X(t) = Z(t) + Y(t)\ , \qquad \text{for each } t \geq 0\ .$$

Since

$$\left[-\inf_{0 \leq u \leq t} \{Z(u)\}\right]^+ \geq [-Z(t)]^+\ , \qquad t \geq 0\ ,$$

$X(t)$ is non-negative for each $t \geq 0$. Suppose that $X(t) > 0$. This implies that

$$-Y(t) = \inf_{0 \leq u \leq t} \{Z(u)\} < 0 \quad \text{and} \quad Z(t) > \inf_{0 \leq u \leq t} \{Z(u)\} ,$$

and hence

$$\inf_{0 \leq u < t} \{Z(u)\} = \inf_{0 \leq u \leq t} \{Z(u)\} .$$

Thus Y does not grow at t, and we have a pair of non-anticipating processes, X and Y, that satisfy (9) and (10).

Suppose now that the process $\{\bar{Y}(t); t \geq 0\}$ is continuous, non-anticipating, non-decreasing from $\bar{Y}(0) = 0$ and grows only if $\bar{X}(t) = 0$, where $\bar{X}(t) = Z(t) + \bar{Y}(t)$. Let τ be the Markov time (with respect to $\{\mathfrak{F}_t; t \geq 0\}$) defined by

$$\tau = \inf\{t \geq 0 : Y(t) \neq \bar{Y}(t)\} ,$$

and suppose that $\tau < +\infty$ and $Y(\tau+) < \bar{Y}(\tau+)$. Then there exists a $\tau' > \tau$ such that $Y(t) < \bar{Y}(t)$ on (τ,τ'). Hence $\bar{Y}$ must grow at each $t \in (\tau,\tau')$ implying that $X(t) < \bar{X}(t) = 0$ on (τ,τ') which contradicts that X is a non-negative process. Similarly, for $\tau < +\infty$ and $Y(\tau+) > \bar{Y}(\tau+)$. It must be then that $\tau = +\infty$, thereby proving the uniqueness of our solution to (9) and (10). That is, the process Y defined by (11) is the unique minimal process needed to maintain the non-negativity of process X as defined by (9). □

Our controlled processes are now defined as follows. If $\lambda = 1$ (reflection), then for each $x \in S$ and each admissible strategy π define process $\{X(t|x, \pi); t \geq 0\}$ exactly as process X in Theorem 1, the notation being enriched to indicate explicitly the dependence of X on π and x. When $\lambda = 0$ (absorbtion), we define $\{X(t|x, \pi); t \geq 0\}$ by $X(t|x, \pi) = X(t \wedge T)$ in Theorem 1 where

$$T = \inf\{t \geq 0 : X(t) = 0\} \quad .$$

In either case (absorbtion or reflection), we call $\{X(t|x, \pi); t \geq 0\}$ the *controlled process generated by strategy* π *and initial state* x.

3.3. Expected Costs and Optimality

Start with a given admissible strategy π and initial control mode $a \in A$. For each pair of equivalence classes k and ℓ in $\{1, 2, \ldots, M\}$, define a counting process $\{Q_{k\ell}(t|a, \pi); t \geq 0\}$ as follows. For $t > 0$, let $Q_{k\ell}(t)|a, \pi)$ be the supremum of all $m \in \{0, 1, 2, \ldots\}$ such that there exist $0 < \tau_1 < \tau_2 < \cdots < \tau_{2m-1} < \tau_{2m} \leq t$, where for $n = 1, 3, 5, \ldots, 2m-1$ we have

(12) $$\pi(\tau_n) \in A_k \; ,$$

(13) $$\pi(u) \in (A_k \cup A_\ell) \qquad \text{for all } u \in (\tau_n, \tau_{n+1})$$

(14) $$\pi(\tau_{n+1}) \in A_\ell \quad .$$

We now define $Q_{k\ell}(0|a, \pi)$ by

$$Q_{k\ell}(0|a, \pi) = \lim_{t \downarrow 0} Q_{k\ell}(t|a, \pi) \quad .$$

(Such a limit exists for all k and ℓ in $\{1, 2, \ldots, M\}$ and all $a \in A$, since $Q_{k\ell}(\cdot|a, \pi)$ is a non-decreasing function of t.)

Clearly, $Q_{k\ell}(t|a, \pi)$ is $\mathfrak{F}_t$ measurable for each $t \geq 0$, all $k, \ell \in \{1, 2, \ldots, M\}$, and all $a \in A$. Note that if for any $t > 0$, there exists $k \in \{1, 2, \ldots, M\}$ and $(\tau, \tau') \subseteq (0,t)$ such that $\theta(u) = k$ on (τ, τ'), then $Q_{kk}(t|a, \pi) = \infty$. But for $t > 0$ and all $k, \ell \in \{1, 2, \ldots, M\}$ where $k \neq \ell$, $Q_{k\ell}(t|a, \pi)$ must be finite by virtue of condition (8) in the definition of admissible strategies. Therefore, except for the case $k = \ell$, we interpret $Q_{k\ell}(t|a, \pi)$ as the number of switches made under strategy π from within equivalence class k to within equivalence class ℓ during the time interval $[0,t]$.

Now let us represent the continuous costs of our system by the function $g : S \times A \to \mathbb{R}$, where

$$g(x,a) = \begin{cases} hx + r_a & \text{if } x > 0 \\ (1-\lambda)\alpha R & \text{if } x = 0 \end{cases}$$

for each $a \in A$. Recalling that $X(\cdot|x, \pi)$ is the controlled process generated by strategy π and initial state x, we define a function V_π on $S \times A$ by

(15) $$V_\pi(x,a) = E\left[\int_0^\infty e^{-\alpha t}[g(X(t|x, \pi), \pi(t))]dt + \sum_{k=1}^{M} \sum_{\ell=1}^{M} \int_0^\infty C_{k\ell}\, e^{-\alpha t}\, dQ^*_{k\ell}(t|a, \pi)\right],$$

where $Q^*_{k\ell}(\cdot|a, \pi)$ is the counting measure associated with counting process $Q_{k\ell}(\cdot|a, \pi)$ and where by convention

(16) $$\int_0^t C_{k\ell}\, dQ^*_{k\ell}(u|a, \pi) = 0 \,, \quad \text{if } C_{k\ell} = 0 \text{ and } Q_{k\ell}(t|a, \pi) = \infty \,.$$

Thus $V_\pi(x,a)$ represents the expected total discounted costs generated by strategy π, given initial state $x \in S$ and initial control mode $a \in A$. We shall call V_π the <u>value function</u> for strategy π.

<u>Theorem 2</u>. Let π be an admissible strategy. Then for all $a \in A$, $|v_\pi(x,a) - hx/\alpha|$ is bounded for all $x \in S$ if and only if $E[\int_0^\infty e^{-\alpha t}\, dQ^*_{k\ell}(t|a, \pi)] < \infty$ for all k and ℓ in $\{1, 2, \ldots, M\}$ where $C_{k\ell} > 0$.

<u>Proof</u>. Fix $(x,a) \in S \times A$. For all $t > 0$, we may bound $g(X(t|x, \pi), \pi(t))$ by $\max\{hX(t|x, \pi) + r^*, (1-\lambda)\, \alpha R\}$ where $r^* = \max\{r_1, r_2, \ldots, r_N\}$. Hence we now set out to bound $E[|\int_0^\infty e^{-\alpha t} g(X(t|x, \pi), \pi(t))dt - hx/\alpha|]$ by effectively doing so with $E[|\int_0^\infty e^{-\alpha t} hX(t|x, \pi)dt - hx/\alpha|]$.

We restrict attention here to the case of reflection, since $E[\int_0^\infty e^{-\alpha t} X(t|x, \pi)dt]$ is greater with a reflected process than with an absorbed one. Let $\{Z_x(t);\ t \geq 0\}$ denote unrestricted Brownian Motion

with drift parameter μ and variance parameter $\sigma^2 > 0$ and initial state x. Define $\{M_x(t);\ t \geq 0\}$ and $\{m_x(t);\ t \geq 0\}$ as

$$M_x(t) = \sup_{0 \leq u \leq t} \{Z_x(u)\} \quad \text{and} \quad m_x(t) = -\inf_{0 \leq u \leq t} \{Z_x(u)\}$$

$$\text{for all} \quad t \geq 0 ,$$

and

$$M_x = \sup_{0 \leq t} \{Z_x(t)\} \quad \text{and} \quad m_x = -\inf_{0 \leq t} \{Z_x(t)\} \ .$$

Letting $\{W_x(t);\ t \geq 0\}$ denote the process Z_x reflected at zero, we have that $W_x(t)$ is equivalent to $x + Z_0(t) + [m_0(t) - x]^+$ for all $t \geq 0$. Therefore we get the following bound for each $t \geq 0$,

$$E[W_x(t)] \leq x + E[Z_0(t) + m_0(t)] \ .$$

Now

$$Z_0(t) + m_0(t) = M_0(t) , \qquad \text{for all} \quad t \geq 0 ,$$

and if $\mu = 0$ we have the exact calculations, e.g., Karlin and Taylor (1975),

$$E[M_0(t)] = \frac{\sigma\sqrt{2t}}{\sqrt{\pi}} \qquad \text{and} \qquad E[W_x(t)] = x + \frac{\sigma\sqrt{2t}}{\sqrt{\pi}} \ .$$

For $\mu < 0$, we note that $M_0(t) \uparrow M_0$ as $t \to \infty$ and that M_0 is exponentially distributed with parameter $2|\mu|/\sigma^2$. Hence

$$E[W_x(t)] \leq x + \frac{\sigma^2}{2|\mu|} , \qquad \text{for all} \quad t \geq 0 \quad \text{and} \quad \mu < 0 \ .$$

For $\mu > 0$, we have that $m_0(t) \uparrow m_0$ as $t \to \infty$ and

$$E[m_0] = \frac{\sigma^2}{2\mu} ,$$

since minus the infinum of unrestricted Brownian Motion with positive drift $|\mu|$, variance σ^2, and initial state zero is equivalent to the supremum of unrestricted Brownian Motion with negative drift $-|\mu|$, variance σ^2, and initial state zero. Thus

$$E[W_x(t)] \leq x + \mu t + \frac{\sigma^2}{2\mu} , \qquad \text{for all } t \geq 0 \text{ and } \mu > 0 .$$

Returning to our reflected controlled process, we have shown that

$$E[X(t|x, \pi)] \leq x + \max\left\{\frac{\sqrt{2+\sigma^*}}{\sqrt{\pi}}, \mu^* t + \frac{\sigma^*}{2\mu_*}\right\} , \qquad \text{for all } t \geq 0 ,$$

where $\sigma^* = \max\{\sigma_1^2, \sigma_2^2, \ldots, \sigma_N^2\}$, $\mu_* = \min\{\min\{|\mu_i| : \mu_i \neq 0,\ i = 1,2,\ldots,N\}, +\infty\}$, and μ^* as before. Thereby in integrating, we can bound $E[\int_0^\infty e^{-\alpha t} X(t|x, \pi)dt]$ by a finite valued linear function of x, namely,

$$E\left[\int_0^\infty e^{-\alpha t} X(t|x, \pi)dt\right] \leq \frac{x}{\alpha} + \frac{\mu^*}{\alpha^2} + \frac{\sqrt{\sigma^*}}{\alpha\sqrt{2\alpha}} + \frac{\sigma^*}{2\alpha\mu_*} .$$

This means that $E[|\int_0^\infty e^{-\alpha t} hX(t|x, \pi) - hx/\alpha|]$ and hence, $E[|\int_0^\infty e^{-\alpha t} g(X(t|x, \pi), \pi(t))dt - hx/\alpha|]$ are each bounded for all $x \in S$. It then follows that $|V_\pi(x,a) - hx/\alpha|$ is likewise bounded in S, if and only if $E[\sum_{k=1}^{M} \sum_{\ell=1}^{M} \int_0^\infty C_{k\ell}\, e^{-\alpha t}\, dQ_{k\ell}^*(t|a, \pi)]$ is. So our

proposed condition on the counting processes $Q_{k\ell}(\cdot|a, \pi)$ is necessary and sufficient for the value function $V_\pi(X,a)$ to be finite for all $(x,a) \in S\times A$, in which case $|V_\pi(x,a)|$ is bounded by a linear function in x for all $a \in A$. □

The <u>optimal value function</u> $V_* : S\times A \to \mathbb{R}$ is defined by

$$(16) \qquad V_*(x,a) = \inf_\pi V_\pi(x,a) , \qquad \text{for all } (x,a) \in S\times A ,$$

where the infinum is taken over all admissible strategies π. Admissible strategy π is called (x,a)-<u>optimal</u> if

$$(17) \qquad V_\pi(x,a) = V_*(x,a) ,$$

and given initial state $x \in S$ and initial control mode $a \in A$, our control problem is to construct an admissible strategy that is (x,a) optimal.

3.4. Stationary Policies

We now define stationary policies and show how each such policy generates an admissible strategy for each initial state x and initial mode $a \in A$.

Definition. A stationary policy is a function $f: S \times A \to A$ satisfying

(18) for each $a \in A$, $f(x,a)$ has finitely many discontinuities in $x \in S$,

(19) if $\lambda = 0$, then $f(0,a) = a$ for each $a \in A$,

(20) for each $(x,a) \in S \times A$, there exists an $\epsilon > 0$ such that $f(y, f(x,a)) = f(x,a)$ for all $y \in (x-\epsilon, x]$ or for all $y \in [x, x+\epsilon)$, unless $\lambda = 0$ and $x = 0$ (see (19)),

(21) for each $a \in A$, the class continuation set $I_a = \{x \in S : f(x,a) \in \theta(a)\}$ is an open subset of S,

(22) if y is a closed boundary point of action continuation set $I^a = \{x \in S : f(x,a) = a\}$ for some $a \in A$, then there exists $\bar{a} \in \theta(a)$ and $\epsilon > 0$ such that $f(x,\bar{a}) = \bar{a}$ for all $x \in \{x \in S : |y-x| < \epsilon$ and $x \notin I^a\}$ and $f(y,\bar{a}) = a$, where $\bar{a} = \lim_{\substack{s \to y \\ s \notin I^a}} f(s,a)$.

Interpret f as a rule for selecting actions through time, the action selected at time t being $f(x,a)$ if the state of the system and control mode in use at that time are x and a, respectively. Note that this rule for selection of actions does not depend on time, and thus the name stationary policy.

For each $a \in A$, the set I^a defined in (22) denotes those states for which f will continue with control mode a. We call I^a the <u>action continuation set</u> associated with mode a under policy f. For each $a \in A$, the set I_a defined in (21) represents those states for which f will continue with an action from equivalence class $\theta(a)$. That is, given mode a in the k-th equivalence class, the action $f(x,a)$ selected by f at state x will also belong to A_k if x is in I_a. Hence we term I_a as the <u>class continuation set</u> associated with mode a under policy f. From conditions (18) through (22), we see that for each $a \in A$, set I_a is the union of a finite number of intervals open in S, and action continuation set I^a is contained in class continuation set I_a.

For example, suppose that $\lambda = 1$ (reflection at zero) and we have three control modes such that $K_{12} = K_{21} = K_{22} = K_{33} = 0$, $K_{13} = K_{23} > 0$, and $K_{31} = K_{32} > 0$. We then have two equivalence classes, $A_1 = \{1,2\}$ and $A_2 = \{3\}$, and equivalence class switching costs $C_{12} = K_{13} = K_{23}$, $C_{21} = K_{31} = K_{32}$, and $C_{11} = C_{22} = 0$. In Figure 1 we illustrate a stationary policy for this particular data. This stationary policy is given by

$$f(x,1) = \begin{cases} 1 & \text{if } x \in [0,s_1] \\ 2 & \text{if } x \in (s_1,s_2] \\ 1 & \text{if } x \in (s_2,s_3) \\ 3 & \text{if } x \in [s_3,s_4] \\ 2 & \text{if } x \in (s_4,\infty) \end{cases} ,$$

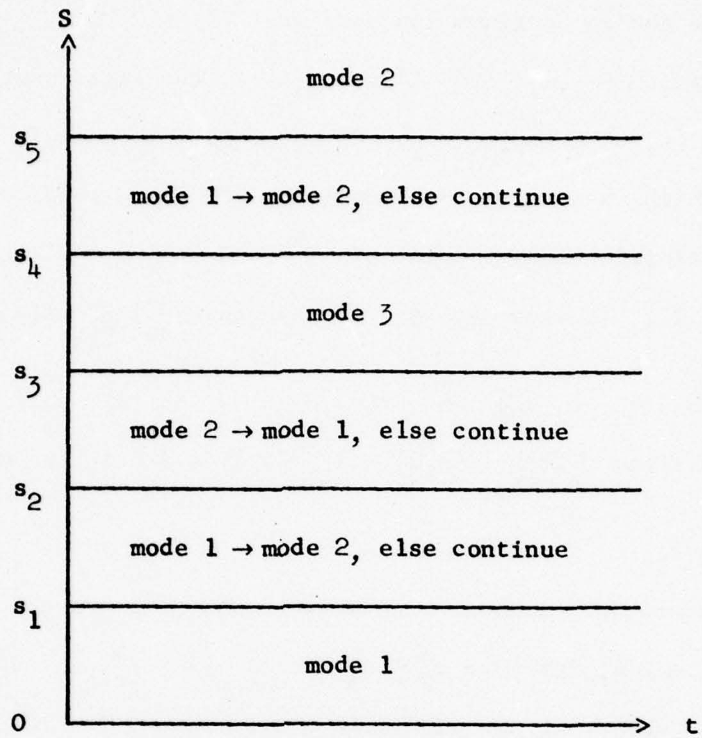

FIGURE 1. An Illustrative Stationary Policy

$$f(x,2) = \begin{cases} 1 & \text{if } x \in [0,s_1] \\ 2 & \text{if } x \in (s_1,s_2] \\ 1 & \text{if } x \in (s_2,s_3) \\ 3 & \text{if } x \in [s_3,s_4] \\ 2 & \text{if } x \in (s_4,\infty) \end{cases} ,$$

and

$$f(x,3) = \begin{cases} 1 & \text{if } x \in [0,s_1] \\ 3 & \text{if } x \in (s_1,s_5) \\ 2 & \text{if } x \in [s_5,\infty) \end{cases} .$$

The action continuation sets are $I^1 = [0,s_1] \cup (s_2,s_3)$, $I^2 = (s_1,s_2] \cup (s_4,\infty)$ and $I^3 = (s_1,s_5)$. The class continuation sets are $I_1 = [0,s_3) \cup (s_4,\infty)$, $I_2 = [0,s_3) \cup (s_4,\infty)$ and $I_3 = (s_1,s_5)$, each of which is open in S. Condition (22) then needs to be verified at s_1 (closed boundary point of I^1) and at s_2 (closed boundary point of I^2). Indeed, at s_1 we have that $\lim_{\substack{s \to s_1 \\ s \notin I^1}} f(s,1) = 2$, $f(x,2) = 2$ on $(s_1,s_2]$ and $f(s_1,2) = 1$. At s_2, (22) is likewise satisfied since $\lim_{\substack{s \to s_2 \\ s \notin I^2}} f(s,2) = 1$, $f(s,1) = 1$ on (s_2,s_3) and $f(s_2,1) = 2$.

We now need some more notation concerning a stationary policy f. For each $a \in A$, let $0 = s_0^a \leq s_1^a < s_2^a < \cdots < s_{n(a)}^a < s_{n(a)+1}^a = \infty$ be the finite number $n(a)$ of discontinuity points in $f(x,a)$. And for all $p = 0, 1, 2, \ldots, n(a)$, denote by b_p^a the constant value of $f(\cdot,a)$ on the interval between s_p^a and s_{p+1}^a, where we have not specified as to whether the endpoints of the interval are open or closed.

The following two results will allow us to define the admissible strategies generated by a stationary policy f, and to characterize the corresponding values as the solutions of certain differential equations.

Theorem 3. Let f be a stationary policy and fix $(x,a) \in S \times A$. Then there exists a unique admissible strategy π such that

$$\theta(\pi(0+)) = \theta(f(x,a)) , \tag{23}$$

(24) $\pi(0+) = f(x,a)$ except for x a closed boundary point of I^a,

and for all $t > 0$

(25) $$\theta(\pi(t+)) = \theta(f(X(t|x, \pi), \pi(t))) ,$$

(26) $\pi(t+) = f(X(t|x, \pi), \pi(t))$ except for $X(t|x, \pi)$ a boundary point of $I^{\pi(t)}$,

where $X(\cdot|x, \pi)$ is the controlled process generated by strategy π and initial state x.

__Definition__. We call π, characterized by (23) through (26), the admissible strategy corresponding to stationary policy f, initial state x and initial control mode a.

__Proof__. Let $a_1 = f(x,a)$ and $p \in \{0, 1, \ldots, n(a_1)\}$ be such that x falls in the interval between $s_p^{a_1}$ and $s_{p+1}^{a_1}$, and $b_p^{a_1} = a_1$. Define $\{\xi(t); t \geq 0\}$ to be the unique reflected Brownian Motion process

$$\xi(t) = \psi(t) + \left[- \inf_{0 \leq u \leq t} \{\psi(u)\}\right]^+ , \qquad t \geq 0 ,$$

where

$$\psi(t) = x + \mu_{a_1} t + \sigma_{a_1} B(t) , \qquad \text{for all } t \geq 0 .$$

Let τ_1 be the first hitting time of s_p^a or s_{p+1}^a by process ξ, and note that τ_1 is a $\{\mathfrak{F}_t;\ t \geq 0\}$ measurable stopping time. We now begin to construct process $\{W(t);\ t > 0\}$ and process $\{Z(t);\ t \geq 0\}$ by defining $W(t) = a_1$ for all $t \in (0, \tau_1]$ and $Z(t) = \xi(t)$ for all $t \in [0, \tau_1]$.

Suppose that $\xi(\tau_1) = s_p^{a_1}$ and that $s_p^{a_1}$ is not a boundary point of I_{a_1}. Thus $s_p^{a_1}$ is a boundary point of I^{a_1}, and we shall first consider the case where $s_p^{a_1} \in I^{a_1}$. We have in $s_p^{a_1}$, then, a closed boundary point of I^{a_1}, and by (22) there exists $a_2 \in \theta(a_1)$ and $f(s_{p+1}^{a_2}, a_2) = a_1$. Now let $\{\psi(t);\ t \geq \tau_1\}$ be the unique process from Theorem 3 in Chapter 2 satisfying

$$(27) \qquad \psi(t) = s_p^{a_1} + \int_{\tau_1}^{t} [\mu_{a_2} X\{\psi(u) < s_p^{a_1}\} + \mu_{a_1} X\{\psi(u) \geq s_p^{a_1}\}]du$$

$$+ \int_{\tau_1}^{t} [\sigma_{a_2} X\{\psi(u) < s_p^{a_1}\} + \sigma_{a_1} X\{\psi(u) \geq s_p^{a_1}\}]\ dB(u)\ ,$$

$$t \geq \tau_1\ ,$$

and let $\{\bar{\xi}(t);\ t \geq \tau_1\}$ be the unique reflected process

$$(28) \qquad \bar{\xi}(t) = \psi(t) + \left[-\inf_{\tau_1 \leq u \leq t} \{\psi(u)\}\right]^+ , \qquad t \geq \tau_1\ .$$

Define τ_2 as the first hitting time of $s_{p+1}^{a_1}$ or $s_{\bar{p}}^{a_2}$ by process $\bar{\xi}$, and note that $\tau_2 > \tau_1$ a.s. and that τ_2 is $\{\mathfrak{F}_t;\ t \geq \tau_1\}$ measurable. Thus define $W(t) = a_2\ X\{\bar{\xi}(t) < s_p^{a_1}\} + a_1\ X\{\bar{\xi}(t) \geq s_p^{a_1}\}$ on $(\tau_1, \tau_2]$ and $Z(t) = \bar{\xi}(t)$ on $(\tau_1, \tau_2]$.

If $\xi(\tau_1) = s_p^{a_1}$, $s_p^{a_1}$ is not a boundary point I_{a_1} but is a boundary point of I^{a_1}, and $s_p^{a_1} \notin I^{a_1}$, then there exists $a_2 \in \theta(a_1)$ and $\bar{p} \in \{0, 1, \ldots, n(a_2)\}$ such that $s_{\bar{p}+1}^{a_2} = s_p^{a_1}$, $b_{\bar{p}}^{a_2} = a_2$ and $f(s_p^{a_1}, a_1) = a_2$. Subsequently we would alter the definitions of $\psi(t)$ and $W(t)$ in the above paragraph to

$$(27') \qquad \psi(t) = s_p^{a_1} + \int_{\tau_1}^{t} \left[\mu_{a_2} \chi\{\psi(u) \le s_p^{a_1}\} + \mu_{a_1} \chi\{\psi(u) > s_p^{a_1}\}\right] dU$$

$$+ \int_{\tau_1}^{t} \left[\sigma_{a_2} \chi\{\psi(u) \le s_p^{a_1}\} + \sigma_{a_1} \chi\{\psi(u) > s_p^{a_1}\}\right] dB(u) \;,$$

$$t \ge \tau_1 \;,$$

and $W(t) = a_2 \chi\{\bar{\xi}(t) \le s_p^{a_1}\} + a_1 \chi\{\bar{\xi}(t) > s_p^{a_1}\}$ on $(\tau_1, \tau_2]$. All other definitions would be the same.

Suppose instead that $\xi(\tau_1) = s_p^{a_1}$ and that $s_p^{a_1}$ is a boundary point of I_{a_1}. Let $a_2 = f(s_p^{a_1}, a_1)$. By virtue of condition (21), there then exists $\bar{p} \in \{0, 1, \ldots, n(a_2)\}$ such that $s_p^{a_1}$ falls in the interior of the interval between $s_{\bar{p}}^{a_2}$ and $s_{\bar{p}+1}^{a_2}$ and $b_{\bar{p}}^{a_2} = a_2$, or such that $s_p^{a_1} = s_{\bar{p}+1}^{a_2}$, $b_{\bar{p}}^{a_2} = a_2$ and $b_{\bar{p}+1}^{a_2} \in \theta(a_2)$.

If $s_{\bar{p}}^{a_2} < s_p^{a_1} < s_{\bar{p}+1}^{a_2}$, we proceed as in the first paragraph of this proof. That is, define $\{\bar{\xi}(t);\ t \ge \tau_1\}$ to be the reflected process

$$(29) \qquad \bar{\xi}(t) = \psi(t) + \left[- \inf_{\tau_1 \le u \le t} \{\psi(u)\}\right]^{+} , \qquad t \ge \tau_1 \;,$$

where

(30) $$\psi(t) = s_p^{a_1} + \mu_{a_2}(t-\tau_1) + \sigma_{a_2}(B(t) - B(\tau_1)) , \qquad t \geq \tau_1 .$$

Letting τ_2 be the first hitting time of $s_{\bar{p}}^{a_2}$ or $s_{\bar{p}+1}^{a_2}$ by process $\bar{\xi}$, we then define processes W and Z by $W(t) = a_2$ on $(\tau_1, \tau_2]$ and $Z(t) = \bar{\xi}(t)$ on $(\tau_1, \tau_2]$. If $s_p^{a_1} = s_{\bar{p}+1}^{a_2}$, we proceed as in the above paragraph. That is, letting $a_3 = b_{\bar{p}+1}^{a_2}$ and $p^* \in \{0, 1, \ldots, n(a_3)\}$ to be such that $b_{p^*}^{a_3} = a_3$, we define $\{\bar{\xi}(t); t \geq \tau_1\}$ to be the reflected process $\{\psi(t); t \geq \tau_1\}$ where

(31) $$\psi(t) = s_p^{a_1} + \int_{\tau_1}^{t} [\mu_{a_2} \chi\{\psi(u) \leq s_p^{a_1}\} + \mu_{a_3} \chi\{\psi(u) > s_p^{a_1}\}]du$$
$$+ \int_{\tau_1}^{t} [\sigma_{a_2} \chi\{\psi(u) \leq s_p^{a_1}\} + \sigma_{a_3} \chi\{\psi(u) > s_p^{a_1}\}]dB(u) ,$$
$$t \geq \tau_1 .$$

We then define processes W and Z by $W(t) = a_2 \chi\{\bar{\xi}(t) \leq s_p^{a_1}\} + a_3 \chi\{\bar{\xi}(t) > s_p^{a_1}\}$ and $Z(t) = \bar{\xi}(t)$ on $(\tau_1, \tau_2]$, where τ_2 is the first hitting time of $s_{\bar{p}}^{a_2}$ or $s_{p^*+1}^{a_3}$ by process $\bar{\xi}$. In both cases of this paragraph, stopping time τ_2 is such that $\tau_2 > \tau_1$ almost surely.

Similarly, we handle all of the possibilities where $\xi(\tau_1) = s_{p+1}^{a}$. Then in conclusion of the situations depicted above, we let $\tilde{x} = Z(\tau_2)$ and $\tilde{a} = W(\tau_2-)$ and follow the same procedures all over again. That is, if $\tilde{x}$ is a boundary point of $I_{\tilde{a}}$ we return to the case discussed in the previous paragraph, and if $\tilde{x}$ is not a boundary point of $I_{\tilde{a}}$

we return to the second paragraph for $\tilde{x} \in I^{\tilde{a}}$, and to the third paragraph for $\tilde{x} \notin I^{\tilde{a}}$.

Thus in this prescribed manner, we continue to build processes W and Z over the time intervals between the $\{\mathfrak{F}_t;\ t \geq 0\}$ measurable stopping times τ_n and τ_{n+1}, for $n = 1, 2, 3, \ldots$. For each $j \in A$, there is a minimal band width

$$\min_{p \in \{0,1,\ldots,n(j)\}} (s^j_{p+1} - s^j_p)$$

of positive width. Hence for every n the length of the time interval between τ_n and τ_{n+1} is bounded away from zero almost surely. This ensures that $\tau_n \to +\infty$ almost surely as $n \to \infty$, and we have that W is indeed a function from $(0,\infty)$ into the action space A.

Finally, if $\lambda = 1$ (reflection) we define process $\{\pi(t);\ t > 0\}$ as being exactly process $\{W(t);\ t > 0\}$, and process $\{X(t);\ t \geq 0\}$ as being exactly process $\{Z(t);\ t \geq 0\}$. If $\lambda = 0$ (absorbtion), we define $\{\pi(t);\ t > 0\}$ by $\pi(t) = W(t \wedge T)$ for all $t > 0$ and $\{X(t);\ t \geq 0\}$ by $X(t) = Z(t \wedge T)$ for all $t \geq 0$, where $T = \inf\{t \geq 0 : Z(t) = 0\}$.

Clearly $\pi(t)$ is $\mathfrak{F}_t$ measurable for all $t > 0$ and since $(\tau_{n+1} - \tau_n)$ is bounded away from zero for every $n = 1, 2, 3, \ldots$, the function $\theta^* : (0,\infty) \to \{1,2,3, \ldots, M\}$ define by $\theta^*(t) = \theta(\pi(t))$ has (a.s.) finitely many discontinuities in each finite interval of time. Therefore π is an admissible strategy. By construction our process Z is the solution process to Theorem 1 of this chapter for this π, so process X is uniquely the controlled process generated by

strategy π and initial state x. Moreover processes π and X uniquely satisfy conditions (23) through (26), which is to say that processes π and $X(\cdot|x, \pi)$ uniquely satisfy (23) - (26), as desired. □

Remark. The fact that $(\tau_{n+1} - \tau_n)$ in the proof above is bounded away from zero for every $n = 1, 2, 3, \ldots$ implies that there exists a constant $U < \infty$ such that $E[Q_{k\ell}(t|a, \pi)] < Ut$, for all $t \geq 0$ and all $k,\ell \in \{1, 2, \ldots, M\}$ where $k \neq \ell$.

Theorem 4. Let f be a stationary policy and for each $(x,a) \in S\times A$, let $\pi(x,a)$ denote the admissible strategy uniquely corresponding to f, x and a. Let $V_f : S\times A \to \mathbb{R}$ be defined as $V_f(x,a) = V_{\pi(x,a)}(x,a)$ for each $(x,a) \in S\times A$. Then V_f is the unique function $V : S\times A \to \mathbb{R}$ to satisfy the following for each $a \in A$:

(32) $$V(\cdot,a) \in C^2_*(S) ,$$

(33) $|V(x,a) - \frac{hx}{\alpha}|$ is bounded for all $x \in S$,

(34) $D_a V(x,a) - \alpha V(x,a) + g(x,a) = 0$ for all x in the interior of I^a, where D_a denotes the differential operator $D_a = \mu_a \frac{\partial}{\partial x} + \frac{1}{2}\sigma_a^2 \frac{\partial^2}{\partial x^2}$,

(35) $V(x,a) = K_{a,f(x,a)} + V(x, f(x,a))$ for all $x \notin I^a$,

(36) $$\lambda V'(0,a) - (1-\lambda) V(0,a) + (1-\lambda)R = 0 .$$

Proof. We begin by showing that V_f satisfies conditions (32) and (36). From the note following Theorem 3 we can see that $E[\int_0^\infty e^{-\alpha t}\, dQ^*_{k\ell}(t|a,\pi)] < \infty$, for every $(x,a) \in S\times A$ and all distinct k and ℓ in $\{1, 2, \ldots, M\}$. Hence by Theorem 2, V_f satisfies (33) for each $a \in A$.

Now fix $a \in A$, and note that I^a is the union of a finite number of intervals in S where the endpoints of each interval are possibly open or closed. Consider $x \in I^a$ with s^a_p and s^a_{p+1} being the endpoints of the interval of I^a containing x. Let $\{Z_a(t);\ t \geq 0\}$ be the Brownian Motion starting in state x with drift parameter μ_a, variance parameter σ_a and absorbtion at boundaries s^a_p and s^a_{p+1}. Associate with process Z_a the same linear holding costs, operational costs, and switching costs as with $X(\cdot|x, \pi)$, and include in this cost structure absorbtion costs $K_{a, f(s^a_p, a)} + V_f(s^a_p, f(s^a_p, a))$ and $K_{a, f(s^a_{p+1}, a)} + V_f(s^a_{p+1}, f(s^a_{p+1}, f(s^a_{p+1}, a))$ at boundary points s^a_p and s^a_{p+1}, respectively. Let $V_a(x)$ denote the conditional expectation of the total discounted cost generated by Z_a in this setting. Define function $F_a : [0,T] \times [s^a_p, s^a_{p+1}] \to \mathbb{R}$ by

$$F_a(t,x) = e^{-\alpha t}\, V_a(x), \qquad (t,x) \in [0,T] \times [s^a_p, s^a_{p+1}] \quad ,$$

where $T = \inf\{t \geq 0 : Z_a(t) \notin (s^a_p, s^a_{p+1})\}$. Now since

$$(37) \qquad V_a(x) = E\Big[\int_0^T e^{-\alpha t}[hZ_a(t) + r_a]dt + e^{-\alpha T}[K_{a, f(Z_a(T), a)} + V_f(Z_a(T), f(Z_a(T), a)]\Big]$$

and Z_a is continuous, we have that $V_a \in C^2([s_p^a, s_{p+1}^a])$. Hence F_a satisfies the conditions of Theorem 2 in Chapter 2, (our extension of Itô's lemma), and the desired result here is that

$$(38) \qquad e^{-\alpha T} V_a(Z_a(T)) = V_a(x) + \int_0^T [-\alpha e^{-\alpha t} V_a(Z_a(t)) + \frac{1}{2} e^{-\alpha t} V_a''(Z_a(t)) \sigma_a^2 + e^{-\alpha t} V_a'(Z_a(t)) \mu_a]dt + \int_0^T e^{-\alpha t} V_a'(Z_a(t)) \sigma_a \, dB(t) \quad .$$

The fact that $V_a \in C^2([s_p^a, s_{p+1}^a])$ results in

$$\int_0^t e^{-2\alpha u}[V_a'(Z_a(u))]^2 \sigma_a^2 \, du < \infty \qquad \text{for all } t \in [0,T] \ ,$$

thereby leading to

$$E\left[\int_0^T e^{-\alpha t} V_a'(Z_a(t)) \sigma_a \, dB(t)\right] = 0 \ .$$

We therefore take expectations in (38) and see that

$$(39) \qquad E[e^{-\alpha T} V_a(Z_a(T))] = V_a(x) + E\left[\int_0^T e^{-\alpha t}[D_a V_a(Z_a(t)) - \alpha V_a(Z_a(t))]dt\right] .$$

But the expectation on the left hand side of (39) is exactly

$$E\left[e^{-\alpha T}[K_{a,f(Z_a(T),a)} + V_a(Z_a(T), f(Z_a(T), a))]\right] ,$$

so combining (37) and (39) we get that

$$E\left[\int_0^T e^{-\alpha t}[\alpha V_a(Z_a(t)) - D_a V_a(Z_a(t))]dt\right] = E\left[\int_0^t e^{-\alpha t}[hZ_a(t) + r_a]dt\right] .$$

This means that $D_a V_a(x) - \alpha V_a(x) + g(x,a) = 0$ for all $x \in [s_p^a, s_{p+1}^a]$, and since $V_a(x) = V_f(x,a)$ on this interval of S, we have that V_π satisfies condition (34).

Now suppose that $x \notin I^a$. Then it follows from the definition of value $V_f(x,a)$ and Theorem 3 that

$$V_f(x,a) = K_{a,f(x,a)} + V_f(x, f(x,a)) ,$$

and so V_F also satisfies (35). Furthermore if x is a continuity point of $f(\cdot, f(x,a))$, then $y \in I^{f(x,a)}$ for all y in some open interval about x. Therefore, as discussed in the previous paragraph, $V_f'(y, f(x,a))$ and $V_f''(y, f(x,a))$ exist for all y in this interval, where V_f' and V_f'' denote the first and second partial derivatives with respect to the first argument of V_f. Hence $V_f'(x,a)$ and $V_f''(x,a)$ both exist.

We have thus far shown $V_f'(x,a)$ and $V_f''(x,a)$ to exist for all x in the interior of I^a and all x in the interior of $I^{f(x,a)}$. This means that $V_f'(x,a)$ and $V_f''(x,a)$ also exist if x is a boundary point of $I_{\theta(a)}$. Suppose now that x is a closed boundary point of I^a. Again let s_p^a and s_{p+1}^a be the endpoints of the interval of I^a containing x, and assume that $x = s_p^a$. Then there exists $j \in \theta(a)$ and $\bar{p} \in \{0, 1, 2, \ldots, n(j)\}$ such that $s_{\bar{p}+1}^j = s_p^a$, $b_{\bar{p}}^j = j$ and

$f(s^j_{p+1}, j) = a$. Following the proof of Theorem 3, let $\{\psi(t);\ t \geq 0\}$ be the unique process defined for $y \in (s^j_p, s^a_{p+1})$ as

$$\psi(t) = y + \int_0^t [\mu_j\ \chi\{\psi(u) < s^a_p\} + \mu_a\ \chi\{\psi(u) \geq s^a_p\}]du$$

$$+ \int_0^t [\sigma_j\ \chi\{\psi(u) < s^a_p\} + \sigma_a\ \chi\{\psi(u) \geq s^a_p\}]dB(u) \quad ,$$

and $\{\xi(t);\ t \geq 0\}$ as the process ψ appropriately absorbed or reflected at zero. Impose on process ξ the linear holding costs, operational costs, and switching costs associated with process $X(\cdot|x, \pi)$, and absorbtion upon hitting s^j_p or s^a_{p+1} at costs $K_{j,f(s^j_p,j)} + V_f(s^j_p, f(s^j_p, j))$ and $K_{a,f(s^a_{p+1},a)} + V_f(s^a_{p+1}, f(s^a_{p+1},a))$, respectively. Let $V_{aj}(y)$ denote the conditional expectation of the total discounted cost generated by ξ in this setting, and we find that $V_{aj}(y)$ is continuously differentiable with respect to y on (s^j_p, s^a_{p+1}) and has a second derivative except at the points s^j_p, s^a_p and s^a_{p+1}. Therefore since $V_{aj}(s^a_p) = V_f(x,a)$, we have that $V'_f(x,a)$ exists, though not necessarily so $V''_f(x,a)$. Similarly if $x = s^a_{p+1}$, and we have shown that $V'_f(x,a)$ exists for all $x \in S$ and $V''_f(x,a)$ exists for all x that are not discontinuity points of $f(x,a)$. Hence V_f satisfies (32).

Finally, in the case of absorbtion property (19) of a stationary policy implies that $V_f(0,a) = R$, thereby validating (36) for $\lambda = 0$. In the case of reflection we have two possibilities to consider, that of state zero included in set I^a and that of zero lying outside of I^a. If $0 \in I^a$, then $V'_f(0,a) = 0$ by application of (34) to the interval between s^a_0 and s^a_1. If $0 \notin I^a$, then $y \in I^{f(0,a)}$ for all y in

some interval $[0, s_1^{f(0,a)})$, and thus, $V_f'(0,a) = V_f'(0, f(0,a)) = 0$ where the second equality follows by the argument of the previous sentence. So (36) holds true also for $\lambda = 1$, and we conclude that V_f is indeed a solution to conditions (32) through (36).

Suppose now that function $V : S \times A \to \mathbb{R}$ also satisfies (32) - (36). Then letting $\Delta(x,a) = V(x,a) - V_f(x,a)$ for all (x,a) in $S \times A$, we would find that the function Δ satisfies these conditions for all $a \in A$:

(33') $\quad |\Delta(x,a) - \frac{2hx}{\alpha}|$, is bounded in S,

(34') $\quad D_a \Delta(x,a) - \alpha\Delta(x,a) = 0$, for all $x \in I^a$,

(35') $\quad \Delta(x,a) = \Delta(x, f(x,a))$, for all $x \notin I^a$,

and

(36') $\quad \lambda\Delta'(0,a) - (1-\lambda)\Delta(0,a) = 0$.

A solution to the second order differential equation (34') would imply that

$$\Delta(x,a) = \gamma_1 e^{\beta_1 x} + \gamma_2 e^{\beta_2 x}, \qquad \text{for all } x \in I^a,$$

where β_1 is the positive root and β_2 the negative root to the quadratic equation $\mu_a\beta + \frac{1}{2}\sigma_a^2\beta^2 - \alpha = 0$. But in order to maintain the required bound (33') and boundary conditions (35') and (36'), it must be that $\gamma_1 = \gamma_2 = 0$. Hence we have that $\Delta(x,a) = 0$ for all $(x,a) \in S \times A$, and V_f is the unique solution to conditions (32) through (36). □

To emphasize the special nature of stationary policies within the class of all admissible strategies we call V_f, as characterized by Theorem 4 above, the <u>return function</u> for stationary policy f. Thus stationary policy f is (x,a)-optimal if $V_f(x,a) = V_*(x,a)$. Finally we conclude this chapter with the concept of everywhere optionality by further calling policy f <u>optimal</u> if it is (x,a)-optimal for all $(x,a) \in S \times A$.

CHAPTER 4

OPTIMAL STATIONARY POLICIES

In this chapter we derive a necessary and sufficient condition for a given stationary policy to be optimal for the control problem formulated in Chapter 3. It will not be shown that there always exists an optimal stationary policy. We conjecture, however, that this is true, and in Chapter 5 we will explicitly produce a stationary policy that is optimal for the general problem when there are two available control modes $N = 2$.

4.1. Optimality of Stationary Policies

We will first prove a preliminary proposition. The main theorem will then be stated and proved.

Proposition 1. Suppose that $V : S{\times}A \to \mathbb{R}$ satisfies the following for all $a \in A$:

(1) $$V(\cdot,a) \in C^2_*(S) ,$$

(2) $$\left|V(x,a) - \frac{hx}{\alpha}\right| , \qquad \text{is bounded for all } x \in S ,$$

(3) $$V(x,a) \leq K_{aj} + V(x,j) , \qquad \text{for all } x \in S \text{ and all } j \in A ,$$

(4) $D_j V(x,a) - \alpha V(x,a) + g(x,j) \geq 0$, for all $x \in S\backslash\{0\}$ and all $j \in \theta(a)$, where we further define differential operator D_j by $D_j V(x,a) = \mu_j V'(x,a) + \frac{1}{2}\sigma_j^2[(V''(x-,a) + V''(x+,a))/2]$ for those points x such that $V''(x,a)$ does not exist,

(5) $$\lambda V'(0,a) - (1-\lambda) V(0,a) + (1-\lambda)R = 0 .$$

Then $V(x,a) \leq V_*(x,a)$ for all $(x,a) \in S \times A$.

<u>Proof.</u> We first note that if function $V : S \times A \to \mathbb{R}$ satisfies (3), then for each $a \in A$, $V(x,j) = V(x,a)$ for all $x \in S$ and all $j \in \theta(a)$. Therefore, we may restate conditions (1) through (5) as

(1') $\bar{V}(\cdot,k) \in C_*^2(S)$, $k \in \{1, 2, \ldots, M\}$,

(2') $|\bar{V}(\cdot,k) - \frac{hx}{\alpha}|$, is bounded for all $x \in S$, $k \in \{1,2,\ldots,M\}$,

(3') $\bar{V}(x,k) \leq C_{k\ell} + \bar{V}(x,\ell)$, for all $x \in S$ and all $k,\ell \in \{1,2,\ldots,M\}$,

(4') $D_j \bar{V}(x,k) - \alpha\bar{V}(x,k) + g(x,j) \geq 0$, for all $x \in S$ and all $j \in A_k$, $k \in \{1, 2, \ldots, M\}$,

(5') $\lambda\bar{V}'(0,k) - (1-\lambda) \bar{V}(0,k) + (1-\lambda)R = 0$, $k \in \{1, 2, \ldots, M\}$,

where for each $(x,k) \in S \times \{1, 2, \ldots, M\}$, we define $\bar{V}(x,k)$ as the common value of $V(x,j)$ for all $j \in A_k$.

Fix $(x,a) \in S \times A$ and let π be an arbitrary admissible strategy. The function $\bar{\theta} : [0,\infty) \to \{1, 2, \ldots, M\}$ defined by

$$\bar{\theta}(t) = \begin{cases} \theta(a) , & \text{if } t = 0 \\ \theta(\pi(t)) , & \text{if } t > 0 \end{cases}$$

has finitely many discontinuities in each finite interval of time. So letting $T_1, T_2, \ldots, T_n, T_{n+1}, \ldots$ denote the discontinuity points of $\bar{\theta}$, we have that $0 = T_0 \leq T_1 < T_2 < \cdots < T_n < T_{n+1} < \cdots$ and $T_n \to +\infty$ almost surely.

Now by inspection of definition (15) in Chapter 3 we see that

$$(6) \qquad V_{\pi}(x,a) = E\left[\sum_{n=0}^{\infty} \left[\int_{T_n}^{T_{n+1}} e^{-\alpha t} g(X(t|x, \pi), \pi(t))dt + e^{-\alpha T_{n+1}} C_{\bar{\theta}(T_{n+1}-),\bar{\theta}(T_{n+1}+)}\right]\right] .$$

Application of condition (4') leads us to

$$(7) \qquad V_{\pi}(x,a) \geq E\left[\sum_{n=0}^{\infty} \left[\int_{T_n}^{T_{n+1}} e^{-\alpha t}[\alpha\bar{V}(x(t|x, \pi), \bar{\theta}(t)) - D_{\pi(t)} \bar{V}(X(t|x, \pi), \bar{\theta}(t))]dt + e^{-\alpha T_{n+1}} C_{\bar{\theta}(T_{n+1}-)+\bar{\theta}(T_{n+1}+)}\right]\right]$$

since $\theta(\pi(t))$ is constant on (T_n, T_{n+1}) for every $n = 0, 1, 2, \ldots$.

Fix $n \in \{0, 1, 2, \ldots\}$. For all $t \in [T_n, T_{n+1}]$ we have that

$$X(t|x, \pi) = X(T_n|x, \pi) + \int_{T_n}^{t} \mu_{\pi(u)}\, du + \int_{T_n}^{t} \sigma_{\pi(u)}\, dB(u) + Y(t)$$

where $X(\cdot|x, \pi)$ and Y uniquely satisfy Theorem 1 in Chapter 3.

Define function $F_n : (T_n, T_{n+1}) \times S \to \mathbb{R}$ by

$$F_n(t,x) = e^{-\alpha t}\, \bar{V}(x, \bar{\theta}(t)) \qquad \text{for all } (t,x) \in (T_n, T_{n+1}) \times S .$$

By condition (1'), $\bar{V}(\cdot, \bar{\theta}(t)) \in C_*^2(S)$ for all $t \in (T_n, T_{n+1})$. Thus F_n satisfies the conditions of Theorem 2 in Chapter 2 and the result here is as follows:

$$(8) \quad e^{-\alpha t}\, \bar{V}(X(t|x, \pi), \bar{\theta}(t))$$

$$= e^{-\alpha T_n}\, \bar{V}(X(T_n|x, \pi), \bar{\theta}(T_n+))$$

$$+ \int_{T_n}^{t} \Big[-\alpha e^{-\alpha u}\, \bar{V}(X(u|x, \pi), \bar{\theta}(u)) + \frac{1}{2} e^{-\alpha u}\, \bar{V}''(X(u|x, \pi), \bar{\theta}(u))\, \sigma^2_{\pi(u)}$$

$$+ e^{-\alpha u}\, \bar{V}'(X(u|x, \pi), \bar{\theta}(u))\, \mu_{\pi(u)} \Big] du$$

$$+ \int_{T_n}^{t} e^{-\alpha u}\, \bar{V}'(X(u|x, \pi), \bar{\theta}(u))\, \sigma_{\pi(u)}\, dB(u)$$

$$+ \int_{T_n}^{t} e^{-\alpha u}\, \bar{V}'(X(u|x, \pi), \bar{\theta}(u))\, dY(u) , \qquad \text{for all } t \in (T_n, T_{n+1}) .$$

Equivalently,

$$(9)\quad e^{-\alpha t}\,\bar{V}(X(t|x,\pi),\bar{\theta}(t))$$

$$= e^{-\alpha T_n}\,\bar{V}(X(T_n|x,\pi),\bar{\theta}(T_n+))$$

$$+ \int_{T_n}^{t} e^{-\alpha u}\Big[D_{\pi(u)}\,\bar{V}(X(u|x,\pi),\bar{\theta}(u)) - \alpha\bar{V}(X(u|x,\pi),\bar{\theta}(u))\Big]\,du$$

$$+ \int_{T_n}^{t} e^{-\alpha u}\,\bar{V}'(X(u|x,\pi),\bar{\theta}(u))\sigma_{\pi(u)}\,dB(u)$$

$$+ \int_{T_n}^{t} e^{-\alpha u}\,\bar{V}'(X(u|x,\pi),\bar{\theta}(u))\,dY(u)\,, \quad \text{for all} \quad t \in (T_n, T_{n+1})\,.$$

If $\lambda = 0$, then the process Y is every where zero, and thus the last integral on the right hand side of (9) equals zero. If $\lambda = 1$, then the non-negative process Y grows only where $X(\cdot|x,\pi)$ is zero. In this case the same integral becomes

$$\int_{T_n}^{t} e^{-\alpha u}\,\bar{V}'(0,\bar{\theta}(u))\,\chi\{X(u|x,\pi)=0\}\,dY(u)$$

and likewise disappears, since $\lambda = 1$ implies that $\bar{V}'(0,\bar{\theta}(u)) = 0$ for all $u \in (T_n, T_{n+1})$ by virtue of condition (5').

Conditions (1') and (2'), the continuity of $X(\cdot|x,\pi)$, and the boundedness of $\sigma_{\pi(\cdot)}$ guarantee that

$$\int_{T_n}^{t} e^{-2\alpha u}\,E\Big[[\bar{V}'(X(u|x,\pi),\bar{\theta}(u))\,\sigma_{\pi(u)}]^2\Big]\,du < \infty \qquad \text{on } (T_n, T_{n+1})\,,$$

which means that

$$E\left[\int_{T_n}^{t} e^{-\alpha u}\, \bar{V}'(X(u|x,\pi), \bar{\theta}(u))\, \sigma_{\pi(u)}\, dB(u)\right] = 0 \quad .$$

Therefore after taking expectations, (9) becomes

$$(10) \quad E\left[e^{-\alpha t}\, \bar{V}(X(t|x,\pi), \bar{\theta}(t))\right]$$

$$= E\left[e^{-\alpha T_n}\, \bar{V}(X(T_n|x,\pi), \bar{\theta}(T_n+))\right]$$

$$+ E\left[\int_{T_n}^{t} e^{-\alpha u}[D_{\pi(u)}\, \bar{V}(X(u|x,\pi), \bar{\theta}(u)) - \alpha\bar{V}(X(u|x,\pi), \bar{\theta}(u))]\right] du$$

$$\text{for all} \quad t \in (T_n, T_{n+1}) \quad .$$

Substituting (10) into (7) for each $n = 0, 1, 2, \ldots$ we get

$$(11) \quad V_{\pi}(x,a) \geq E\left[\sum_{n=0}^{\infty}\left[e^{-\alpha T_n}\, \bar{V}(X(T_n|x,\pi), \bar{\theta}(T_n+))\right.\right.$$

$$- e^{-\alpha T_{n+1}}\, \bar{V}(X(T_{n+1}|x,\pi), \bar{\theta}(T_{n+1}-))$$

$$\left.\left. + e^{-\alpha T_{n+1}}\, C_{\bar{\theta}(T_{n+1}-), \bar{\theta}(T_{n+1}+)}\right]\right]$$

$$= E\left[\bar{V}(x, \theta(a)) + \sum_{n=1}^{\infty}\left[e^{-\alpha T_n}\, \bar{V}(X(T_n|x,\pi), \bar{\theta}(T_n+))\right.\right.$$

$$\left.\left. - e^{-\alpha T_n}\, \bar{V}(X(T_n|x,\pi), \bar{\theta}(T_n-)) + e^{-\alpha T_n}\, C_{\bar{\theta}(T_n-), \bar{\theta}(T_n+)}\right]\right] .$$

Finally, condition (3') implies that

$$\bar{V}(X(T_n|x, \pi), \bar{\theta}(T_n-)) \le C_{\bar{\theta}(T_n-),\bar{\theta}(T_n+)} + \bar{V}(X(T_n|x, \pi), \bar{\theta}(T_n+)) ,$$

for each n ,

so that we conclude that

$$(12) \quad V_\pi(x,a) \ge E\Big[\bar{V}(x, \theta(a)) + \sum_{n=1}^{\infty} e^{-\alpha T_n} \Big[\bar{V}(X(T_n|x, \pi), \bar{\theta}(T_n+)) - \Big[C_{\bar{\theta}(T_n-),\bar{\theta}(T_n+)} + \bar{V}(X(T_n|x, \pi), \bar{\theta}(T_n+))\Big] + C_{\bar{\theta}(T_n-),\bar{\theta}(T_n+)}\Big]\Big]$$

$$= \bar{V}(x, \theta(a)) .$$

And since $\bar{V}(x, \theta(a)) = V(x,a)$, we have as desired that $V_\pi(x,a) \ge V(x,a)$ and $V_*(x,a) \ge V(x,a)$. □

<u>Theorem 1</u>. Let f be a fixed stationary policy and denote by $V_f(x,a)$ its return function on $S \times A$. Then f is optimal if and only if the following three conditions are satisfied for each $a \in A$:

$$(13) \qquad V_f(x,a) = \min_{j \in A} \{K_{aj} + V_f(x,j)\} , \qquad \text{for all } x \in S ,$$

(14) $$\min_{j \in L(x,a)} \{D_j V_f(x,j) - \alpha V_f(x,j) + g(x,j)\} = 0 \quad \text{for all } x \in S\backslash\{0\}$$

where $L(x,a) = \{j \in A : V_f(x,a) = K_{aj} + V_f(x,j)\}$ and D_j is as before, and

(15) $$\lambda V_f'(0,a) - (1-\lambda)\, V_f(0,a) + (1-\lambda)R = 0 \quad .$$

Proof. From Theorem 4 in Chapter 3 we have that $V_f(\cdot,a) \in C_*^2(S)$ for each $a \in A$ and that $|V_f(x,a) - hx/\alpha|$ is bounded in $S \times A$. Since $\theta(a) \subseteq L(x,a)$ for all $(x,a) \in S \times A$, condition (14) here implies condition (4) in Proposition 1. Hence all of the conditions in Proposition 1 are satisfied, which immediately results in $V_f(x,a) = V_*(x,a)$ everywhere on $S \times A$.

Suppose now that stationary policy f is optimal. Condition (15) is exactly the boundary condition satisfied by all stationary policies, so it holds for an optimal policy.

If (13) fails, then there exists $x \in S$ and a and j in A such that $V_f(x,a) > K_{aj} + V_f(x,j)$. Then since $V_f(\cdot,a)$ is continuous on S, there exists an $\epsilon > 0$ such that the following hold

(16) $$V_f(y,a) > K_{aj} + V_f(y,j) \; , \quad \text{for all } y \in [x-\epsilon, x+\epsilon] \; ,$$

and

(17) $$\left[K_{aj} + \frac{hx}{\alpha} + \frac{r_j}{\alpha} + \frac{\mu_j}{\alpha^2}\right]\left[1 - E[e^{-\alpha T}]\right] - \frac{\mu_j}{\alpha} E[Te^{-\alpha T}] + E\left[e^{-\alpha T} V_f(\psi(T), a)\right] \leq V_f(x,a) \quad ,$$

where $\{\psi(t);\ t \geq 0\}$ is the process, $\psi(t) = x + \mu_j t + \sigma_j B(t)$ and T is the stopping time $T = \inf\{t \geq 0 : \psi(t) \notin [x-\epsilon, x+\epsilon]\}$. Now define the admissible strategy π by

(18) $\pi(t) = j$, for $t \in (0,T]$

(19) $\theta(\pi(t+)) = \theta(f(X(t|x, \pi), \pi(t)))$, for $t > T$,

(20) $\pi(t+) = f(X(t|x, \pi), \pi(t))$, for $t > T$ except for $X(t|x, \pi)$ a boundary point of $I^{\pi(t)}$,

where $X(\cdot|x, \pi)$ is the control process generated by π and x. That is, strategy π uses action j until time T and then follows policy f every afterward. Hence

$$V_\pi(x,a) = K_{aj} + E\Big[\int_0^t e^{-\alpha t} g(X(t|x, \pi), \pi(t))dt$$

$$+ e^{-\alpha T} K_{j,f(X(T(x,\pi),j)} + e^{-\alpha T} V_f(X(T|x,\pi), f(X(T|x,\pi),j))\Big]$$

$$= K_{aj} + E\Big[\int_0^T e^{-\alpha t}[h\psi(t) + r_j]dt\Big] + E[e^{-\alpha T} V_f(\psi(T), j)] ,$$

and using (16) and (17) we have that

$$V_\pi(x,a) < K_{aj} + E\Big[\int_0^T e^{-\alpha t}[h\psi(t) + r_j]dt\Big] + E\Big[e^{-\alpha T}[V_f(\psi(T),a)-K_{aj}]\Big] \leq V_f(x,a).$$

This, however, contradicts the optimality of f proving that (13) must hold.

If (14) fails, then there exists $x \in S\backslash\{0\}$ and a and j in A such that $V_f(x,a) = K_{aj} + V_f(x,j)$ and $D_j V_f(x,j) - \alpha V_f(x,j) + g(x,j) < 0$. Since $V_f(\cdot,j) \in C_*^2(S)$, there exists an $\epsilon > 0$ such that

$$D_j V_f(y,j) - \alpha V_f(y,j) + g(y,j) < 0, \quad \text{for all } y \in [x-\epsilon, x+\epsilon],$$

and we define process ψ, stopping time T, and admissible strategy π exactly as in the previous paragraph. Now,

$$\begin{aligned} V_\pi(x,a) &= K_{aj} + E\left[\int_0^T e^{-\alpha t} g(\psi(t), j)dt\right] + E[e^{-\alpha T} V_f(\psi(T),j)] \\ &< K_{aj} + E\left[\int_0^T e^{-\alpha t}[\alpha V_f(\psi(t),j) - D_j V_f(\psi(t),j)]dt\right] \\ &\quad + E[e^{-\alpha T} V_f(\psi(T),j)] \ . \end{aligned}$$

But as we saw in the proof of Proposition 1, we can apply our extension of Itô's lemma (Theorem 2 in Chapter 2) to get that

$$E[e^{-\alpha T} V_f(\psi(T),j)] = V_f(x,j) + E\left[\int_0^T e^{-\alpha t}[D_j V_f(\psi(t), j) - \alpha V_f(\psi(t),j)]dt\right] .$$

Hence

$$V_\pi(x,a) < K_{aj} + V_f(x,j) = V_f(x,a) \ ,$$

and we have again contradicted the optimality of f, thereby concluding that (14) is also necessary for stationary policy f to be optimal. □

4.2. An Interpretation of the Optimality Conditions

We would now like to interpret the optimality conditions of Theorem 1 and the proof given there. Let $V_f : S \times A \to \mathbb{R}$ be the return function corresponding to an optimal stationary policy f. Assume an initial starting state x and an initial control mode $a \in A$.

Suppose that you, as the controller, switch immediately to mode j at time zero and continue to use j over the interval $[0,t]$, and thereafter follow policy f. Letting $U_{aj}(x,t)$ denote your expected total discounted cost we have

$$(21) \qquad U_{aj}(x,t) = K_{aj} + E\left[\int_0^t e^{-\alpha u}[hX_j(u) + r_j]du + e^{-\alpha t} V_f(X_j(t),j)\right],$$

where $\{X_j(t);\ t \geq 0\}$ is the Brownian Motion process starting in state x with drift parameter μ_j and variance parameter σ_j, properly reflected or absorbed at zero. Using Theorem 2 in Chapter 2 (Itô's lemma) to evaluate $E[V_f(X_j(t), j)]$ in (21) above, we can approximate $U_{aj}(x,t)$ by

$$K_{aj} + hxt + r_j t + (1-\alpha t)\ [V_f(x,j) + tD_j V_f(x,j)] + o(t)\ ,$$

since

$$E\left[\int_0^t e^{-\alpha u}[hX_j(u) + r_j]du\right] = hxt + r_j t + o(t)$$

and

$$E\left[e^{-\alpha t} V_f(X_j(t), j)\right] = [1 - \alpha t + o(t)]\ [V_f(x,j) + tD_j V_f(x,j) + o(t)] \ .$$

Therefore,

$$U_{aj}(x,t) = [K_{aj} + V_f(x,j)] + [D_j V_f(x,j) - \alpha V_f(x,j) + hx + r_j]t + o(t),$$

and for $V_f(x,a)$ to be the optimal return it must be that

$$V_f(x,a) = \min_{j \in A} U_{aj}(x,t)$$

for all sufficiently small t.

Hence we can summarize our optimality conditions (13) and (14) of Theorem 1, by demanding that $V_f(x,a)$ satisfy the single condition

$$(22) \qquad \min_{j \in A} \{[K_{aj} + V_f(x,j) - V_f(x,a)] + t[D_j V_f(x,j) - \alpha V_f(x,j) + g(x,j)]\} = 0 \ ,$$

for all small enough t .

We call (22) a <u>lexicographic</u> <u>minimum</u> <u>condition</u> since it requires first that

$$(23) \qquad V_f(x,a) = \min_{j \in A} \{K_{aj} + V_f(x,j)\} \ ,$$

and then that

$$(24) \qquad \min_{j \in L(x,a)} \{D_j V_f(x,j) - \alpha V_f(x,j) + g(x,j)\} = 0$$

where $L(x,a)$ is the set of $j \in A$ that achieve the minimum in (23). Equation (22) is the Bellman equation of dynamic programming, specialized to our control problem.

CHAPTER 5

SOME EXPLICIT SOLUTIONS

In this chapter we present some explicit solutions to the optimal control problems formulated in Chapter 3. That is, using the necessary and sufficient optimality conditions developed in Chapter 4 we construct stationary policies that are optimal. We deal with the system that has two available control modes, i.e., $N = 2$ and $A = \{1,2\}$. The two control modes are characterized by the drift and variance parameter pairs (μ_1, σ_1^2) and (μ_2, σ_2^2), and we have labelled the modes 1 and 2 so that $\mu_1 \geq \mu_2$.

In the first section we deal with an absorbing barrier at the boundary and the particular cost structure of linear holding cost rate $h = 0$, operational cost rates $r_1 = r_2 = 0$, boundary cost $R \neq 0$, and switching costs $K_{12} = K_{21} = 0$. We call this form of our control problem a death penalty problem, and we show the simplest of stationary policies, one that forever uses the same control mode regardless of initial state and initial mode, to be optimal. When R is positive, it will be optimal to always use mode 1 if and only if μ_1, σ_1^2, μ_2, σ_2^2 and interest rate α satisfy specified relationships. However, when R is negative those same parameter combinations lead to the optimality of the policy that always uses mode 2.

Section 5.2 treats the absorbing barrier problem with zero switching costs and general cost parameters h, r_1, r_2 and R. We immediately prove this optimal control problem to be equivalent to one in which there is absorbtion, no switching costs, no linear holding costs, one zero operational

cost, and one non-negative operational cost. We then determine an optimal stationary policy that selects actions as a function of state only and that is characterized by a single critical number $z \in S$. One control mode is used whenever the state of the system is above level z, and the other mode is used when the state is below the critical level. For certain realizations of the cost and diffusion parameters, z will be zero and our optimal policy will be one that simply uses the same control mode everywhere on S. For other parameter situations, the critical number will be positive and explicitly stated as the unique solution to a complicated transcendental equation.

In Section 5.3 we present an optimal stationary policy and optimal return function for the control problem with no switching costs and reflection at the boundary. With a reflecting barrier the controller must be concerned with controlling linear holding costs, and to avoid the computational complexities seen in Section 5.2, we will assume zero operational costs. We will show that if $\sigma_1^2 \geq \sigma_2^2$, then an optimal policy is to always use mode 2. If $\sigma_1^2 < \sigma_2^2$, we have what we call a <u>tortoise-hare problem</u>. (Mode 1 is the "tortoise" and mode 2 is the "hare".) In this case our optimal policy is also a single critical number policy, and again the critical number z is the unique positive solution to a transcendental equation.

In Section 5.4 we add to the tortoise-hare problem of Section 5.3 a positive symmetric switching cost $K = K_{12} = K_{21} > 0$. With such switching costs we show that there exists an optimal stationary policy that is a function of both current state and current mode. This optimal

policy involves two critical numbers z and Z, where $0 \leq z < Z \leq \infty$. Control mode 2 is used whenever the state of the system is above Z and mode 1 is used whenever the state falls short of z. When the state is inbetween the critical numbers, the controller maintains the control mode currently in use. As in the case of zero switching costs, the critical numbers are characterized by complicated formulas.

5.1. A Death Penalty Problem

Suppose that we have a two mode system, absorbtion at zero, and only a non-zero cost R at the boundary. That is, the linear holding cost rate h, operational cost rates r_1 and r_2, and switching costs K_{12} and K_{21} are all zero. Let π be any admissible strategy and given initial state $x \in S$, define the random variables

$$T_x(y) = \inf\{t \geq 0 : X(t|x, \pi) \leq x-y\}, \quad \text{for all } y \in [0,x] .$$

Corresponding to π then is the value function

$$(1) \qquad V_\pi(x,a) = R\, E\left[e^{-\alpha T_x(x)}\right], \qquad \text{for all } (x,a) \in S \times A .$$

Let us first consider the stationary policy f_1 of always using control mode 1. Since the switching costs are zero here, the return function associated with f_2 is such that $V_{f_1}(x,1) = V_{f_1}(x,2)$ for all $x \in S$. Therefore we can suppress the second argument in V_{f_1} and represent the return function as a function $V_{f_1} : S \to \mathbb{R}$ of initial state only.

For $(x,a) \in S \times A$ let $\pi(x,a)$ be the unique admissible strategy corresponding to f_1, x and a. Using the strong Markov property of $X(\cdot|x, \pi(x,a))$ we then have that

$$E\left[\exp(-\alpha(T_x(y) + T_x(x) - T_x(y))); T_x(y) < \infty\right]$$

$$= E\left[\exp(-\alpha T_x(y)); T_x(y) < \infty\right] E\left[\exp(-\alpha(T_x(x) - T_x(y))); T_x(y) < \infty\right],$$

$$\text{for all } y \in [0,x] .$$

Now since $X(\cdot|x, \pi(x,a))$ has stationary and independent increments, the random variable $[T_x(x) - T_x(y)]$ has the same distribution as the random variable $T_y(y)$. Therefore

$$E[\exp(-\alpha T_x(x))] = E[\exp(-\alpha T_x(y))]\, E[\exp(-\alpha T_y(y))], \quad \text{for all } y \in [0,x].$$

Thus we have the exponential function

$$\text{(2)} \qquad V_{f_1}(x) = RE[\exp(-\alpha T_x(x))] = Re^{-\beta x}, \qquad \text{for all } x \in S ,$$

where β is a real-valued function of the parameters and is non-negative to insure that the expectation in (2) decreases as x increases on S. Substituting (2) into Theorem 4 of Chapter 3 we see that

$$Re^{-\beta x}[-\mu_1\beta + \frac{1}{2}\sigma_1^2\beta^2 - \alpha] = 0 , \qquad \text{for all } x \in S ,$$

and hence β must solve the quadratic equation

$$\alpha + \mu_1\beta - \frac{1}{2}\sigma_1^2\beta^2 = 0 .$$

Since we desire β to be non-negative, it must be that

$$\beta = \frac{\mu_1 + \sqrt{\mu_1^2 + 2\alpha\sigma_1^2}}{\sigma_1^2} ,$$

and we have completely determined the return function for policy f_1.

Proceeding to evaluate the necessary and sufficient optimality conditions, we have

$$D_1 V_{f_1}(x) - \alpha V_{f_1}(x) = R\, e^{-\beta x}[-\mu_1\beta + \frac{1}{2}\sigma_1^2\beta^2 - \alpha] = 0 , \quad \text{for all } x \in S,$$

and

$$D_2 V_{f_1}(x) - \alpha V_{f_1}(x) = R\, e^{-\beta x}[-\mu_2\beta + \frac{1}{2}\sigma_2^2\beta^2 - \alpha] , \quad \text{for all } x \in S,$$

and $V_{f_1}(0) = R$. Therefore V_{f_1} will be the optimal return function and f_1 an optimal stationary policy if and only if

$$R[\frac{1}{2}\sigma_2^2\beta^2 - \mu_2\beta - \alpha] \geq 0 . \tag{3}$$

Condition (3) can be shown to be equivalent to

$$R\left[(\mu_1 + \sqrt{\mu_1^2 + 2\alpha\sigma_1^2})\,(\mu_1\sigma_2^2 - \mu_2\sigma_1^2) + \alpha\sigma_1^2(\sigma_2^2 - \sigma_1^2)\right] \geq 0 , \tag{4}$$

and further inspection leads us to conclude that under any of the following combinations of diffusion parameters and interest rate it will always be optimal to use control mode 1:

$$[1] \quad \begin{cases} R > 0 \\ \sigma_1^2 \leq \sigma_2^2 , \end{cases}$$

$$[2] \quad \begin{cases} R > 0 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq \mu_2 > 0 \\ \dfrac{\mu_1}{\mu_2} > \dfrac{\sigma_1^2}{\sigma_2^2} \\ \alpha \leq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2 - \sigma_2^2)^2} \end{cases}$$

$$[3] \quad \begin{cases} R > 0 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq 0 \geq \mu_2 \\ \mu_1 \neq \mu_2 \\ \alpha \leq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2 - \sigma_2^2)^2} , \end{cases}$$

$$[4] \quad \begin{cases} R > 0 \\ \sigma_1^2 > \sigma_2^2 \\ 0 > \mu_1 \geq \mu_2 \\ \alpha \leq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2 - \sigma_2^2)^2} \end{cases}$$

$$
[5] \quad \begin{cases} R < 0 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq \mu_2 > 0 \\ \mu_1\sigma_2^2 - \mu_2\sigma_1^2 \leq 0 \, , \end{cases}
$$

$$
[6] \quad \begin{cases} R < 0 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq \mu_2 > 0 \\ \dfrac{\mu_1}{\mu_2} > \dfrac{\sigma_1^2}{\sigma_2^2} \\ \alpha \geq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1 - \mu_2)}{(\sigma_1^2 - \sigma_2^2)^2} \end{cases}
$$

$$
[7] \quad \begin{cases} R < 0 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq 0 \geq \mu_2 \\ \mu_1 \neq \mu_2 \\ \alpha \geq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1 - \mu_2)}{(\sigma_1^2 - \sigma_2^2)^2} \end{cases}
$$

$$
[8] \quad \begin{cases} R < 0 \\ \sigma_1^2 > \sigma_2^2 \\ 0 > \mu_1 \geq \mu_2 \\ \alpha \geq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1 - \mu_2)}{(\sigma_1^2 - \sigma_2^2)^2} \end{cases}
$$

Next consider the policy f_2 of always using control mode 2. Similar analysis results in a return function V_{f_2} that is a function of state only and in the following

$$V_{f_2}(x) = Re^{-\rho x} , \qquad \text{for all } x \in S ,$$

where

$$\rho = \frac{\mu_2 + \sqrt{\mu_2^2 + 2\alpha\sigma_2^2}}{\sigma_2^2} ,$$

$$D_1 V_{f_2}(x) - \alpha V_{f_2}(x) = Re^{-\rho x}[-\mu_1\rho + \tfrac{1}{2}\sigma_1^2\rho^2 - \alpha] , \qquad \text{for all } x \in S ,$$

$$D_2 V_{f_2}(x) - \alpha V_{f_2}(x) = Re^{-\rho x}[-\mu_2\rho + \tfrac{1}{2}\sigma_2^2\rho^2 - \alpha] = 0, \qquad \text{for all } x \in S ,$$

and $V_{f_2}(0) = R$. Therefore V_{f_2} is the optimal return function if and only if

$$(5) \qquad R\left[(\mu_2 + \sqrt{\mu_2^2 + 2\alpha\sigma_2^2})\,(\mu_2\sigma_1^2 - \mu_1\sigma_2^2) + \alpha\sigma_2^2(\sigma_1^2 - \sigma_2^2)\right] \geq 0 .$$

We thus find it optimal to use mode 2 always whenever one of the following combinations arises:

$$[9] \qquad \begin{cases} R < 0 \\ \sigma_1^2 \leq \sigma_2^2 \end{cases}$$

$$[10]\quad \begin{cases} R < 0 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq \mu_2 > 0 \\ \dfrac{\mu_1}{\mu_2} > \dfrac{\sigma_1^2}{\sigma_2^2} \\ \alpha \leq \dfrac{2(\mu_2\sigma_1^2 - \mu_1\sigma_2^2)\ (\mu_2-\mu_1)}{(\sigma_2^2-\sigma_1^2)^2} \end{cases} ,$$

$$[11]\quad \begin{cases} R < 0 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq 0 \geq \mu_2 \\ \mu_1 \neq \mu_2 \\ \alpha \leq \dfrac{2(\mu_2\sigma_1^2 - \mu_1\sigma_2^2)\ (\mu_2-\mu_1)}{(\sigma_2^2-\sigma_1^2)^2} \end{cases} ,$$

$$[12]\quad \begin{cases} R < 0 \\ \sigma_1^2 > \sigma_2^2 \\ 0 > \mu_1 \geq \mu_2 \\ \alpha \leq \dfrac{2(\mu_2\sigma_1^2 - \mu_1\sigma_2^2)\ (\mu_2-\mu_1)}{(\sigma_2^2-\sigma_1^2)^2} \end{cases} ,$$

$$[13]\quad \begin{cases} R > 0 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq \mu_2 \geq 0 \\ \mu_1\sigma_2^2 - \mu_2\sigma_1^2 \leq 0 , \end{cases}$$

$$[14]\quad \begin{cases} R > 0 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq \mu_2 > 0 \\ \dfrac{\mu_1}{\mu_2} > \dfrac{\sigma_1^2}{\sigma_2^2} \\ \alpha \geq \dfrac{2(\mu_2\sigma_1^2 - \mu_1\sigma_2^2)\ (\mu_2-\mu_1)}{(\sigma_2^2 - \sigma_1^2)^2} \end{cases}$$

$$[15]\quad \begin{cases} R > 0 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq 0 \geq \mu_2 \\ \mu_1 \neq \mu_2 \\ \alpha \geq \dfrac{2(\mu_2\sigma_1^2 - \mu_1\sigma_2^2)\ (\mu_2-\mu_1)}{(\sigma_2^2-\sigma_1^2)^2} \end{cases}$$

$$[16]\quad \begin{cases} R > 0 \\ \sigma_1^2 > \sigma_2^2 \\ 0 > \mu_1 \geq \mu_2 \\ \alpha \geq \dfrac{2(\mu_2\sigma_1^2 - \mu_1\sigma_2^2)\ (\mu_2-\mu_1)}{(\sigma_2^2-\sigma_1^2)^2} \end{cases}$$

The combinations [1] through [16] exhaust all possible values for the diffusion parameters and the positive interest rate. Hence we have completely solved the death penalty problem and have shown that an optimal single band policy always exists. That is, it will be optimal to either always use control mode 1 regardless of initial state and initial mode, or it will be optimal to always use mode 2.

5.2. Absorbtion and No Switching Costs

We treat here the general two control absorbing barrier problem with linear holding cost rate h, operational cost rates r_1 and r_2 for mode 1 and mode 2, respectively, boundary cost R, zero switching costs and positive interest rate α. For a given admissible strategy π we have the following value function

$$V_\pi(x,a) = E\left[\int_0^T e^{-\alpha t}[hX(t|x,\pi) + r_{\pi(t)}]dt + Re^{-\alpha T}\right], \qquad (x,a) \in S\times A ,$$

where T is the time of absorbtion for the controlled process $X(\cdot|x,\pi)$. The next proposition will allow us to rid our cost structure of all linear holding costs.

<u>Proposition 1</u>. Let π be an arbitrary admissible strategy and let $x \in S$. Then

$$E\left[\int_0^T e^{-\alpha t}\, X(t|x,\pi)dt\right] = \frac{x}{\alpha} + E\left[\int_0^T \frac{1}{\alpha}\, e^{-\alpha t}\, \mu_{\pi(t)}\, dt\right]$$

where $T = \inf\{t \geq 0 : X(t|x,\pi) = 0\}$.

<u>Proof</u>. Fix admissible strategy π and state x. Let $\mu : [0,\infty) \to \{0, \mu_1, \mu_2, \ldots, \mu_N\}$ and $\sigma : [0,\infty) \to \{0, \sigma_1, \sigma_2, \ldots, \sigma_N\}$ be the following functions

$$\mu(t) = \begin{cases} \mu_{\pi(t)}, & t \leq T \\ 0, & t > T, \end{cases}$$

and

$$\sigma(t) = \begin{cases} \sigma_{\pi(t)}, & t \leq T \\ 0, & t > T. \end{cases}$$

Now define the following Itô process $\{X(t);\ t \geq 0\}$,

$$X(t) = x + \int_0^t \mu(u)du + \int_0^t \sigma(u)\ dB(u)\ , \qquad t \geq 0\ ,$$

and note that

$$E\left[\int_0^T e^{-\alpha t}\ X(t|x,\ \pi)dt\right] = E\left[\int_0^\infty e^{-\alpha t}\ X(t)dt\right] .$$

Since σ is bounded on $[0,\infty)$ we have that

$$E\left[\int_0^t \sigma(u)\ dB(u)\right] = 0\ , \qquad \text{for all} \quad t \in [0,\infty)\ .$$

Hence

$$E[X(t)] = x + E\left[\int_0^t \mu(u)du\right], \qquad \text{for all} \quad t \in [0,\infty)\ ,$$

and a simple change in the order of integration leads to the desired result. □

So returning to our control problem, let

$$\bar{r} = \min\{\frac{h}{\alpha}\mu_1 + r_1, \frac{h}{\alpha}\mu_2 + r_2\} ,$$

$$\bar{r}_1 = \frac{h}{\alpha}\mu_1 + r_1 - \bar{r} ,$$

$$\bar{r}_2 = \frac{h}{\alpha}\mu_2 + r_2 - \bar{r} ,$$

and

$$\bar{R} = R + \frac{\bar{r}}{\alpha} .$$

By virtue of the above proposition, we then have that

$$V_\pi(,a) = \frac{hx}{\alpha} - \frac{\bar{r}}{\alpha} + E\left[\int_0^T e^{-\alpha t} \bar{r}_{\pi(t)} dt + \bar{R} e^{-\alpha t}\right] , \quad (x,a) \in S \times A ,$$

for any admissible strategy π. Hence minimizing $V_\pi(x,a)$ over all admissible strategies is equivalent to minimizing $E[\int_0^T e^{-\alpha t} \bar{r}_{\pi(t)} dt + \bar{R}e^{-\alpha T}]$, and we have reduced the original problem to one where there are no linear holding costs, one zero operational cost, and one non-negative operational cost.

Let us look first at the case $r_1 = 0$. Assume also that $r_2 > 0$, since otherwise we would have the death penalty problem which has already been solved. Since the switching costs are zero, the return function associated with any stationary policy f is such that $V_f(x,1) = V_f(x,2)$ for all $x \in S$. We therefore, as before, represent the return function V_f as a function of state only.

We begin with the single band policy f_1 of always using mode 1 and recall from Section 5.1 that

$$V_{f_1}(x) = Re^{-\beta x} , \qquad \text{for all } x \in S ,$$

$$\beta = \frac{\mu_1 + \sqrt{\mu_1^2 + 2\alpha\sigma_1^2}}{\sigma_1^2} \quad .$$

Likewise, the single band policy f_2 of always using control mode 2 leads to

$$V_{f_2}(x) = \frac{r_2}{\alpha} + \left[R - \frac{r_2}{\alpha}\right] e^{-\rho x} \,, \qquad \text{for all } x \in S \,,$$

where

$$\rho = \frac{\mu_2 + \sqrt{\mu_2^2 + 2\alpha\sigma_2^2}}{\sigma_2^2} \quad .$$

We then have the following relationships on S

(6) $$D_1 V_{f_1}(x) - \alpha V_{f_1}(x) = 0 \,,$$

(7) $$D_2 V_{f_1}(x) - \alpha V_{f_1}(x) + r_2 = r_2 + \left[\frac{\sigma_2^2}{2}\beta^2 - \mu_2\beta - \alpha\right] Re^{-\beta x} \,,$$

(8) $$V_{f_1}(0) = R \,,$$

(9) $$D_1 V_{f_1}(x) - \alpha V_{f_2}(x) = -r_2 + \left[\frac{\sigma_1^2}{2}\rho^2 - \mu_1\rho - \alpha\right]\left[R - \frac{r_2}{\alpha}\right] e^{-\rho x} \,,$$

(10) $$D_2 V_{f_2}(x) - \alpha V_{f_2}(x) + r_2 = 0 \,,$$

and

(11) $$V_{f_2}(0) = R \,.$$

Let Δ_1 and Δ_2 denote the following two functions of the parameters $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \alpha)$,

$$\Delta_1 = \frac{\sigma_2^2}{2} \beta^2 - \mu_2\beta - \alpha ,$$

and

$$\Delta_2 = \frac{\sigma_1^2}{2} \rho^2 - \mu_1\rho - \alpha .$$

The optimality conditions of Theorem 1, Chapter 4 will then be satisfied for policy f_1 if

$$r_2 + \Delta_1 Re^{-\beta x} \geq 0 , \qquad \text{for all } x \in S , \tag{12}$$

and will be satisfied for f_2 if

$$-r_2 + \Delta_1\left[R - \frac{r_2}{\alpha}\right] e^{-\rho x} \geq 0 , \qquad \text{for all } x \in S . \tag{13}$$

We have seen in evaluation of (4) that $\Delta_1 R \geq 0$ if and only if one of the situations [1] through [8] holds true. Therefore the policy of always using mode 1 is optimal whenever we have one of these following cases:

[17] $\begin{cases} r_2 > 0, \ R \geq 0 \\ \sigma_1^2 \leq \sigma_2^2 , \end{cases}$

$$[18]\quad \begin{cases} r_2 > 0,\ R \geq 0 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq \mu_2 > 0 \\ \dfrac{\mu_1}{\mu_2} > \dfrac{\sigma_1^2}{\sigma_2^2} \\ \alpha \leq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2 - \sigma_2^2)^2} \end{cases} ,$$

$$[19]\quad \begin{cases} r_2 > 0,\ R \geq 0 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq 0 \geq \mu_2 \\ \mu_1 \neq \mu_2 \\ \alpha \leq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2 - \sigma_2^2)^2} \end{cases} ,$$

$$[20]\quad \begin{cases} r_2 > 0,\ R \geq 0 \\ \sigma_1^2 > \sigma_2^2 \\ 0 > \mu_1 \geq \mu_2 \\ \alpha \leq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2 - \sigma_2^2)^2} \end{cases} ,$$

$$[21]\quad \begin{cases} r_2 > 0,\ R \geq 0 \\ -\Delta_1 R \leq r_2 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq \mu_2 \geq 0 \\ \mu_1\sigma_2^2 - \mu_2\sigma_1^2 \leq 0 , \end{cases}$$

$$
[22] \quad \begin{cases} r_2 > 0, \ R \geq 0 \\ -\Delta_1 R \leq r_2 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq \mu_2 > 0 \\ \dfrac{\mu_1}{\mu_2} > \dfrac{\sigma_1^2}{\sigma_2^2} \\ \alpha \geq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2-\sigma_2^2)^2} \end{cases} ,
$$

$$
[23] \quad \begin{cases} r_2 > 0, \ R \geq 0 \\ -\Delta_1 R \leq r_2 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq 0 \geq \mu_2 \\ \mu_1 \neq \mu_2 \\ \alpha \geq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2-\sigma_2^2)^2} \end{cases} ,
$$

$$
[24] \quad \begin{cases} r_2 > 0, \ R \geq 0 \\ -\Delta_1 R \leq r_2 \\ \sigma_1^2 > \sigma_2^2 \\ 0 > \mu_1 \geq \mu_2 \\ \alpha \geq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2-\sigma_2^2)^2} \end{cases} ,
$$

$$
[25] \quad \begin{cases} r_2 > 0, \ R < 0 \\ -\Delta_1 R \leq r_2 \\ \sigma_1^2 \leq \sigma_2^2 \end{cases} ,
$$

$$
[26] \quad \begin{cases} r_2 > 0, \ R < 0 \\ -\Delta_1 R \leq r_2 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq \mu_2 > 0 \\ \dfrac{\mu_1}{\mu_2} > \dfrac{\sigma_1^2}{\sigma_2^2} \\ \alpha \leq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2 - \sigma_2^2)^2} \end{cases}
$$

$$
[27] \quad \begin{cases} r_2 > 0, \ R < 0 \\ -\Delta_1 R \leq r_2 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq 0 \geq \mu_2 \\ \mu_1 \neq \mu_2 \\ \alpha \leq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2 - \sigma_2^2)^2} \end{cases} ,
$$

$$
[28] \quad \begin{cases} r_2 > 0, \ R < 0 \\ -\Delta_1 R \leq r_2 \\ \sigma_1^2 > \sigma_2^2 \\ 0 > \mu_1 \geq \mu_2 \\ \alpha \leq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2 - \sigma_2^2)^2} \end{cases} ,
$$

$$[29]\quad \begin{cases} r_2 > 0,\ R < 0 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq \mu_2 \geq 0 \\ \mu_1\sigma_2^2 - \mu_2\sigma_1^2 \leq 0\ , \end{cases}$$

$$[30]\quad \begin{cases} r_2 > 0,\ R < 0 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq \mu_2 > 0 \\ \dfrac{\mu_1}{\mu_2} > \dfrac{\sigma_1^2}{\sigma_2^2} \\ \alpha \geq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2 - \sigma_2^2)^2}\ , \end{cases}$$

$$[31]\quad \begin{cases} r_2 > 0,\ R < 0 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq 0 \geq \mu_2 \\ \mu_1 \neq \mu_2 \\ \alpha \geq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2 - \sigma_2^2)^2}\ , \end{cases}$$

$$[32]\quad \begin{cases} r_2 > 0,\ R < 0 \\ \sigma_1^2 > \sigma_2^2 \\ 0 > \mu_1 \geq \mu_2 \\ \alpha \geq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2 - \sigma_2^2)^2}\ . \end{cases}$$

Inspection of (13) shows that for $r_2 < 0$ and every set of diffusion parameters

$$\lim_{x \uparrow \infty} \left\{ -r_2 + \Delta_1 \left[R - \frac{r_2}{\alpha} \right] e^{-\rho x} \right\} < 0 \quad .$$

Hence the policy of always using mode 2 can not be optimal. We have then thoroughly investigated the optimality of the single band policies f_1 and f_2. However, the above parameter combinations [17] through [32] do not totally exhaust the possible range of parameters.

So let us next look at the following two band stationary policy f_z

$$f_z(x,1) = f_z(x,2) = \begin{cases} 2 \,, & \text{if } x \in [0,z) \\ 1 \,, & \text{if } x \in [z,\infty) \,. \end{cases}$$

This control policy selects actions according to a function of current state only and is characterized by the single critical number $z \in S$. Whenever the state of the system is z or greater control mode 1 is to be used, and whenever the state of the system is below level z policy f_z selects mode 2.

Again let $\pi(x,a)$ denote the admissible strategy corresponding to policy f_z, initial state x and initial control mode a, and let $T_x(y) = \inf\{t \geq 0 : X(t|x, \pi(x,a)) = y\}$ for all $y \in S$. We then see that the following hold for the return function V_{f_z}

$$(14) \qquad V_{f_z}(x) = E\left[\int_0^{T_x(z)} e^{-\alpha t} r_2 dt + e^{-\alpha T_x(z)} V_{f_z}(z);\ T_x(z) < T_x(0)\right]$$

$$+ E\left[\int_0^{T_x(0)} e^{-\alpha t} r_2 dt + e^{-\alpha T_x(0)} R;\ T_x(0) < T_x(z)\right] ,$$

$$\text{for } x \in [0,z] ,$$

and

$$(15) \qquad V_{f_z}(x) = E\left[e^{-\alpha T_x(z)} V_{f_z}(z)\right] , \qquad \text{for } x \in [z,\infty) .$$

From (15) we get

$$(16) \qquad V_{f_z}(x) = V_{f_z} e^{-\beta(x-z)} , \qquad \text{for } x \geq z$$

since as before we can replace $E[e^{-\alpha T_x(z)}]$ by the exponential function $\exp(-\beta(x-z))$ on $[z,\infty)$.

We now wish to evaluate the two expectations in (14). Define φ and ψ to be the following functions on $[0,z]$

$$\varphi(x) = \begin{cases} 0 , & \text{if } x = 0 \\ E\left[e^{-\alpha T_x(z)} ;\ T_x(z) < T_x(0)\right] , & \text{if } 0 < x < z \\ 1 , & \text{if } x = z , \end{cases}$$

and

$$\psi(x) = \begin{cases} 1 , & \text{if } x = 0 \\ E\left[e^{-\alpha T_x(0)} ;\ T_x(0) < T_x(z)\right] , & \text{if } 0 < x < z \\ 0 , & \text{if } x = z . \end{cases}$$

For $x \in [0,z]$ let $\{X(t);\ t \geq 0\}$ be the Brownian Motion starting in state x with drift parameter μ_2, variance parameter σ_2 and absorbtion at levels zero and z. Associating first with process X zero holding costs, zero operational costs, zero switching costs, and absorbtion costs of zero and one at the origin and level z, respectively, we can see that $\varphi(x)$ represents the expected returns in this setting. If instead the absorbtion costs are one at the origin and zero at level z, then $\psi(x)$ represents the expected total discounted costs. Theorem 4 of Chapter 3 characterizes the return functions associated with stationary policies and so leads us to the condition

$$(17) \qquad D_2\varphi(x) - \alpha\varphi(x) = D_2\psi(x) - \alpha\psi(x) = 0\ , \qquad \text{for all } x \in [0,z]\ .$$

Returning to calculation of (20), let γ be the function on S defined as

$$\gamma(x) = E[e^{-\alpha\bar{T}}]\ , \qquad \text{for all } x \in S\ ,$$

where $\bar{T}$ is the first hitting time of zero by the unrestricted Brownian Motion starting in state x with drift parameter μ_2 and variance parameter σ_2. We have shown before that

$$\gamma(x) = e^{-\rho x}\ , \qquad \text{for all } x \in S\ ,$$

and we note that

(18) $$\gamma(x) = \psi(x) + \varphi(x)\,\gamma(z)\,, \qquad \text{for all } x \in [0,z]\,.$$

For $x \in (-\infty,z]$ let $\{Z(t);\ t \geq 0\}$ be the Brownian Motion starting in state x with drift parameter μ_2, variance parameter σ_2 and absorbtion at level z (but no absorbtion at zero). Then the function ξ defined on $(-\infty,z]$ by

$$\xi(x) = E[e^{-\alpha T^*}]\,,$$

where $T^* = \inf\{t \geq 0 : Z(t) = z\}$, must be an increasing exponential function in x. That is,

$$\xi(x) = e^{-\eta(x-z)}\,, \qquad \text{for } x \in (-\infty,z]\,,$$

where η is a real-valued function of the parameters $(\mu_2, \sigma_2^2, \alpha)$ and where η is non-positive. Assessing only a boundary cost of one to the process Z at the level z, we apply Theorem 4, Chapter 3 to get

$$D_2\xi(x) - \alpha\xi(x) = 0\,, \qquad \text{for all } x \in (-\infty, z]\,.$$

Therefore we want η to be the non-positive solution to the quadratic equation

$$\alpha + \mu_2\eta - \frac{1}{2}\sigma_2^2\,\eta^2 = 0$$

which means that

$$\eta = \frac{\mu_2 - \sqrt{\mu_2^2 + 2\alpha\sigma_2^2}}{\sigma_2^2} \quad .$$

Since

$$(19) \qquad \xi(x) = \varphi(x) + \psi(x)\,\xi(0) \,, \qquad \text{for all } x \in [0,z] \,,$$

we have in (18) and (19) two equations in the two unknowns $\varphi(x)$ and $\psi(x)$. The solutions are as follows

$$\varphi(x) = \frac{e^{-\eta x} - e^{-\rho x}}{e^{-\eta z} - e^{-\rho z}} \,, \qquad \text{for } x \in [0,z] \,,$$

and

$$\psi(x) = e^{-\rho x} - \left[\frac{e^{-\eta x} - e^{-\rho x}}{e^{-\eta z} - e^{-\rho z}}\right] e^{-\rho z} \,, \quad \text{for } x \in [0,z] \,.$$

We can now write (14) as

$$(20) \qquad V_{f_z}(x) = \frac{r_2}{\alpha} + \left[V_{f_z}(z) - \frac{r_2}{\alpha}\right]\left[\frac{e^{-\eta x} - e^{-\rho x}}{e^{-\eta z} - e^{-\rho z}}\right]$$

$$+ \left[R - \frac{r_2}{\alpha}\right]\left[\frac{e^{-\eta z-\rho x} - e^{-\rho z-\eta x}}{e^{-\eta z} - e^{-\rho z}}\right] \,, \qquad \text{for all } x \in [0,z] \,,$$

and there remains only the unknown value $V_{f_z}(z)$ in our explicit solution of the return function associated with policy f_z. Using condition (32) of Theorem 4 in Chapter 3 we can characterize $V_{f_z}(z)$ from (16) and (20) by

$$(21)\quad \left[V_{f_z} - \frac{r_2}{\alpha}\right]\left[\frac{\rho e^{-\rho z} - \eta e^{-\eta z}}{e^{-\eta z} - e^{-\rho z}}\right] + \left[R - \frac{r_2}{\alpha}\right]\left[\frac{\eta e^{-\rho z-\eta z} - e^{-\eta z-\rho z}}{e^{-\eta z} - e^{-\rho z}}\right]$$

$$= -\beta V_{f_z}(z) \quad .$$

We then solve for $V_{f_z}(z)$ to find that

$$(22)\quad V_{f_z}(z) = \frac{r_2}{\alpha}\left[\frac{\rho e^{-\rho z}(1-e^{-\eta z}) - \eta e^{-\eta z}(1-e^{-\rho z})}{e^{-\eta z} - e^{-\rho z}}\right]\left[\frac{\rho e^{-\rho z} - \eta e^{-\eta z}}{(\rho-\beta)\, e^{-\rho z} - (\eta-\beta)\, e^{-\eta z}}\right]$$

$$+ R\left[\frac{(\rho-n)\, e^{-(\eta+\rho)z}}{e^{-\eta z} - e^{-\rho z}}\right]\left[\frac{\rho e^{-\rho z} - \eta e^{-\eta z}}{(\rho-\beta)\, e^{-\rho z} - (\eta-\beta)\, e^{-\eta z}}\right] .$$

Thus we can complete the explicit solution of V_{f_z} on S as follows

$$(23)\quad V_{f_z}(x) = \frac{r_2}{\alpha}\left[(e^{-\eta z} - e^{-\rho z})^2\, [(\rho-\beta)\, e^{-\rho z} - (\eta-\beta)\, e^{-\eta z}]\right]^{-1}$$

$$\cdot \Big\{ e^{-\eta z}(e^{-\eta z}-e^{-\rho z})\, [(\rho-\beta)e^{-\rho z} - (\eta-\beta)e^{-\eta z}]\, (1 - e^{-\eta(x-z)} - e^{-\rho x})$$

$$- e^{-\rho z}(e^{-\eta z}-e^{-\rho z})\, [(\rho-\beta)e^{-\rho z} - (\eta-\beta)e^{-\eta z}](1 - e^{-\rho(x-z)} - e^{-\eta x})$$

$$+ (\rho e^{-\rho z} - \eta e^{-\eta z})\, [\rho e^{-\rho z}(1-e^{-\eta z}) - \eta e^{-\eta z}(1-e^{-\rho z})](e^{-\eta x}-e^{-\rho x})\Big\}$$

$$+ R\left[(e^{-\eta z} - e^{-\rho z})^2\, [(\rho-\beta)e^{-\rho z} - (\eta-\beta)e^{-\eta z}]\right]^{-1}$$

$$\cdot \Big\{ e^{-(\eta+\rho)z}(e^{-\eta z} - e^{-\rho z})[(\rho-\beta)e^{-\rho z} - (\eta-\beta)e^{-\eta z}](e^{-\rho(x-z)} - e^{-\eta(x-z)})$$

$$+ (\rho-\eta)e^{-(\eta+\rho)z}(\rho e^{-\rho z} - \eta e^{-\eta z})(e^{-\eta x}-e^{-\rho x})\Big\} \quad ,$$

for $0 \leq x \leq z$,

and

$$(24) \qquad V_{f_z}(x) = \frac{r_2}{\alpha}\left[(e^{-\eta z} - {}^{-\rho z})\ [(\rho-\beta)e^{-\rho z} - (\eta-\beta)e^{-\eta z}]\right]^{-1}$$

$$\cdot \left\{(\rho e^{-\rho z} - \eta e^{-\eta z})\ [\rho e^{-\rho z}(1-e^{-\eta z}) - \eta e^{-\eta z}(1-e^{-\rho z})]e^{-\beta(x-z)}\right\}$$

$$+ R\left[(e^{-\eta z} - e^{-\rho z})\ [(\rho-\beta)e^{-\rho z} - (\eta-\beta)e^{-\eta z}]\right]^{-1}$$

$$\cdot \left\{(\rho-\eta)e^{-(\eta+\rho)z}(\rho e^{-\rho z} - \eta e^{-\eta z})e^{-\beta(x-z)}\right\} \quad ,$$

$$\text{for} \quad x \geq z \quad .$$

Let us turn now to the necessary and sufficient optimality conditions of Theorem 1, Chapter 4. We immediately verify that

$$D_2 V_{f_z}(x) - \alpha V_{f_z}(x) + r_2 = \left[V_{f_z}(z) - \frac{r_2}{\alpha}\right] [D_2\, \varphi(x) - \alpha\varphi(x)]$$

$$+ \left[R - \frac{r_2}{\alpha}\right] [D_2\psi(x) - \alpha\psi(x)] = 0 \quad ,$$

$$\text{for all} \quad x \in [0,z] \quad ,$$

$$D_1 V_{f_z}(x) - \alpha V_{f_z}(x) = V_{f_z}(x)\ [\tfrac{1}{2}\, \sigma_1^2\beta^2 - \mu_1\beta - \alpha]\ e^{-\beta(x-z)} = 0 \quad ,$$

$$\text{for all} \quad x \in [z,\infty) \quad ,$$

and $V_{f_z}(0) = R$, and we shall judiciously choose our critical number z so as to satisfy the remaining optimality conditions. Those two remaining optimality conditions are concerned with showing the non-negativity of

$D_1 V_{f_z}(x) - \alpha V_{f_z}(x)$ on $[0,z]$ and the non-negativity of $D_2 V_{f_z}(x) - \alpha V_{f_z} + r_2$ on $[z,\infty)$. Letting Δ_3 denote the following function of the parameters $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \alpha)$,

$$\Delta_3 = \frac{\sigma_1^2}{2}\eta^2 - \mu_1\eta - \alpha ,$$

we see that

$$(25)\quad D_1 V_{f_z}(x) - \alpha V_{f_z}(x)$$

$$= \frac{r_2}{\alpha}\Big[(e^{-\eta z} - e^{-\rho z})^2\,[(\rho-\beta)e^{-\rho z} - (\eta-\beta)e^{-\eta z}]\Big]^{-1}$$

$$\cdot \Big\{e^{-\eta z}(e^{-\rho z} - e^{-\eta z})\,[(\rho-\beta)e^{-\rho z} - (\eta-\beta)e^{-\eta z}]\,[\Delta_3 e^{-\eta(x-z)} + \Delta_2 e^{-\rho x} + \alpha]$$

$$+ e^{-\rho z}(e^{-\rho z} - e^{-\eta z})[(\rho-\beta)e^{-\rho z} - (\eta-\beta)e^{-\eta z}][\Delta_2 e^{-\rho(x-z)} + \Delta_3 e^{-\eta x} + \alpha]$$

$$+ (\rho e^{-\rho z} - \eta e^{-\eta z})[\rho e^{-\rho z}(1-e^{-\eta z}) - \eta e^{-\eta z}(1-e^{-\rho z})][\Delta_3 e^{-\eta x} - \Delta_2 e^{-\rho x}]\Big\}$$

$$+ R\Big[(e^{-\eta z} - e^{-\rho z})^2\,[(\rho-\beta)e^{-\rho z} - (\eta-\beta)e^{-\eta z}]\Big]^{-1}$$

$$\cdot \Big\{ e^{-(\eta+\rho)z}(e^{-\eta z} - e^{-\rho z})[(\rho-\beta)e^{-\rho z} - (\eta-\beta)e^{-\eta z}][\Delta_2 e^{-\rho(x-z)} - \Delta_3 e^{-\eta(x-z)}]$$

$$+ (\rho-n)e^{-(\eta+\rho)z}(\rho e^{-\rho z} - \eta e^{-\eta z})[\Delta_3 e^{-\eta x} - \Delta_2 e^{-\rho x}]\Big\} ,$$

$$\text{for } 0 \le x \le z ,$$

and

(26) $D_2 V_{f_z} - \alpha V_{f_z} + r_2$

$$= \frac{r_2}{\alpha} \left[(e^{-\eta z} - e^{-\rho z}) [(\rho-\beta) e^{-\rho z} - (\eta-\beta) e^{-\eta z}] \right]^{-1}$$

$$\cdot \left\{ (\rho e^{-\rho z} - \eta e^{-\eta z}) [\rho e^{-\rho z}(1-e^{-\eta z}) - \eta e^{-\eta z}(1-e^{-\rho z})] \, \Delta_1 e^{-\beta(x-z)} \right\}$$

$$+ R \left[(e^{-\eta z} - e^{-\rho z}) [(\rho-\beta) e^{-\rho z}) - (\eta-\beta) e^{-\eta z}] \right]^{-1}$$

$$\cdot \left\{ (\rho-\eta) e^{-(\eta+\rho) z} (\rho e^{-\rho z} - \eta e^{-\eta z}) \, \Delta_1 \, e^{-\beta(x-z)} \right\} + r_2 \ ,$$

for $x \geq z$.

Thus, defining the function H on S as

(27) $$H(x) = D_2 V_{f_z}(x) - D_1 V_{f_z}(x) + r_2 \ , \qquad \text{for all } x \in S \ ,$$

we wish to find a $z \in S$ such that $H(x) < 0$ for $x \in [0,z)$, $H(z) = 0$, and $H(x) > 0$ for $x \in (z,\infty)$. The absence of switching costs in the optimality conditions leads us to require that $V''_{f_z}(z-) = V''_{f_z}(z+)$, and so we have the following equation

(28) $$\frac{r_2}{\alpha} \beta[\rho(\rho-\beta)e^{-\rho z} - \eta(\eta-\beta)e^{-\eta z}] + (R - \frac{r_2}{\alpha}) (\rho-\eta)(\beta-\rho)(\eta-\beta)e^{-(\eta+\rho) z} = 0 \ .$$

View the left hand side of (28) as a function of $z \in [0,\infty)$, and let $F(z)$ denote its value. In our analysis of $F(z)$ we may restrict attention to the situations

$$(29) \qquad \begin{cases} r_2 > 0, \ R > 0 \\ -\Delta_1 R > r_2 \end{cases}$$

and

$$(30) \qquad \begin{cases} r_2 > 0, \ R < 0 \\ -\Delta_1 R > r_2 \end{cases}$$

since we have shown policy f_1 to be optimal otherwise.

Much tedious manipulation of (27) and (28) under situations (29) and (30) implies separating the remaining parameter combinations into two groups. Group I includes:

$$\begin{cases} r_2 > 0, \ R < 0 \\ -\Delta_1 R > r_2 \\ \mu_2 < 0 \\ \sigma_1^2 \leq \sigma_2^2 \\ \alpha \geq \dfrac{2\mu_2(\mu_2\sigma_1^2 - \mu_1\sigma_2^2)}{\sigma_2^4} \end{cases} ,$$

$$\begin{cases} r_2 > 0, \ R < 0 \\ -\Delta_1 R > r_2 \\ \mu_2 \geq 0 \\ \sigma_1^2 \leq \sigma_2^2 \end{cases} ,$$

$$
\left\{
\begin{array}{l}
r_2 > 0, \; R < 0 \\
-\Delta_1 R > r_2 \\
\sigma_1^2 > \sigma_2^2 \\
\mu_1 \geq \mu_2 > 0 \\
\dfrac{\mu_1}{\mu_2} > \dfrac{\sigma_1^2}{\sigma_2^2} \\
\alpha \leq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\,(\mu_1-\mu_2)}{(\sigma_1^2-\sigma_2^2)^2}
\end{array}
\right. ,
$$

$$
\left\{
\begin{array}{l}
r_2 > 0, \; R < 0 \\
-\Delta_1 R > r_2 \\
\sigma_1^2 > \sigma_2^2 \\
\mu_1 \geq 0 \geq \mu_2 \\
\mu_1 \neq \mu_2 \\
-\mu_2\sigma_1^2(1-2\sigma_2^2) \leq \mu_1\sigma_2^4 \\
\dfrac{2\mu_2(\mu_2\sigma_1^2-\mu_1\sigma_2^2)}{\sigma_2^4} \leq \alpha \leq \dfrac{2(\mu_1\sigma_2^2-\mu_2\sigma_1^2)\,(\mu_1-\mu_2)}{(\sigma_1^2-\sigma_2^2)^2}
\end{array}
\right. ,
$$

$$
\left\{
\begin{array}{l}
r_2 > 0, \; R < 0 \\
-\Delta_1 R > r_2 \\
\sigma_1^2 > \sigma_2^2 \\
0 > \mu_1 \geq \mu_2 \\
-\mu_2\sigma_1^2(1-2\sigma_2^2) \leq \mu_1\sigma_2^4 \\
\dfrac{2\mu_2(\mu_2\sigma_1^2-\mu_1\sigma_2^2)}{\sigma_2^4} \leq \alpha \leq \dfrac{2(\mu_1\sigma_2^2-\mu_2\sigma_1^2)\,(\mu_1-\mu_2)}{(\sigma_1^2-\sigma_2^2)^2}
\end{array}
\right. ,
$$

$$
\begin{cases}
r_2 > 0,\ R > 0 \\
-\Delta_1 R > r_2 \\
\sigma_1^2 > \sigma_2^2 \\
\mu_1 \geq 0 \geq \mu_2 \\
\mu_1 \neq \mu_2 \\
-\mu_2\sigma_1^2(1-2\sigma_2^2) > \mu_1\sigma_2^4 \\
\alpha \geq \dfrac{2\mu_2(\mu_2\sigma_1^2-\mu_1\sigma_2^2)}{\sigma_2^4}
\end{cases}
,
$$

$$
\begin{cases}
r_2 > 0,\ R > 0 \\
-\Delta_1 R > r_2 \\
\sigma_1^2 > \sigma_2^2 \\
\mu_1 \geq 0 \geq \mu_2 \\
\mu_1 \neq \mu_2 \\
-\mu_2\sigma_1^2(1-2\sigma_2^2) \leq \mu_1\sigma_2^4 \\
\alpha \geq \dfrac{2(\mu_1\sigma_2^2-\mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2-\sigma_2^2)^2}
\end{cases}
,
$$

$$
\begin{cases}
r_2 > 0,\ R > 0 \\
-\Delta_1 R > r_2 \\
\sigma_1^2 > \sigma_2^2 \\
0 > \mu_1 \geq \mu_2 \\
-\mu_2\sigma_1^2(1-2\sigma_2^2) > \mu_1\sigma_2^4 \\
\alpha \geq \dfrac{2\mu_2(\mu_2\sigma_1^2-\mu_1\sigma_2^2)}{\sigma_2^4}
\end{cases}
,
$$

and

AD-A065 174   STANFORD UNIV CALIF DEPT OF OPERATIONS RESEARCH   F/G 12/1
SOME PROBLEMS IN THE OPTIMAL CONTROL OF DIFFUSIONS.(U)
DEC 78 D SHENG   N00014-76-C-0418
UNCLASSIFIED   TR-84   NL

2 OF 2
ADA
065174

END
DATE
FILMED
4-79
DDC

$$\begin{cases} r_2 > 0,\ R > 0 \\ -\Delta_1 R > r_2 \\ \sigma_1^2 > \sigma_2^2 \\ 0 > \mu_1 \geq \mu_2 \\ -\mu_2\sigma_1^2(1-2\sigma_2^2) \leq \mu_1\sigma_2^4 \\ \alpha \geq \dfrac{2(\mu_1\sigma_2^2-\mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2-\sigma_2^2)^2} \end{cases} .$$

Group II includes:

$$\begin{cases} r_2 > 0,\ R < 0 \\ -\Delta_1 R > r_2 \\ \mu_2 < 0 \\ \sigma_1^2 \leq \sigma_2^2 \\ \alpha < \dfrac{2\mu_2(\mu_2\sigma_1^2-\mu_1\sigma_2^2)}{\sigma_2^4} \end{cases} ,$$

$$\begin{cases} r_2 > 0,\ R < 0 \\ -\Delta_1 R > r_2 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq 0 \geq \mu_2 \\ \mu_1 \neq \mu_2 \\ -\mu_2\sigma_1^2(1-2\sigma_2^2) \leq \mu_1\sigma_2^4 \\ \alpha \leq \dfrac{2\mu_2(\mu_2\sigma_1^2-\mu_1\sigma_2^2)}{\sigma_2^4} \end{cases} ,$$

$$\begin{cases} r_2 > 0,\ R < 0 \\ -\Delta_1 R > r_2 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq 0 \geq \mu_2 \\ \mu_1 \neq \mu_2 \\ -\mu_2\sigma_1^2(1-2\sigma_2^2) > \mu_1\sigma_2^4 \\ \alpha \leq \dfrac{2(\mu_1\sigma_2^2-\mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2-\sigma_2^2)^2} \end{cases},$$

$$\begin{cases} r_2 > 0,\ R < 0 \\ -\Delta_1 R > r_2 \\ \sigma_1^2 > \sigma_2^2 \\ 0 > \mu_1 \geq \mu_2 \\ -\mu_2\sigma_1^2(1-2\sigma_2^2) \leq \mu_1\sigma_2^4 \\ \alpha \leq \dfrac{2\mu_2(\mu_2\sigma_1^2-\mu_1\sigma_2^2)}{\sigma_2^4} \end{cases},$$

$$\begin{cases} r_2 > 0,\ R < 0 \\ -\Delta_1 R > r_2 \\ \sigma_1^2 > \sigma_2^2 \\ 0 > \mu_1 \geq \mu_2 \\ -\mu_2\sigma_1^2(1-2\sigma_2^2) > \mu_1\sigma_2^4 \\ \alpha \leq \dfrac{2(\mu_1\sigma_2^2-\mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2-\sigma_2^2)^2} \end{cases},$$

$$\begin{cases} r_2 > 0,\ R > 0 \\ -\Delta_1 R > r_2 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq \mu_2 \geq 0 \\ \mu_1\sigma_2^2 - \mu_2\sigma_1^2 \leq 0 \end{cases} ,$$

$$\begin{cases} r_2 > 0,\ R > 0 \\ -\Delta_1 R > r_2 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq \mu_2 > 0 \\ \dfrac{\mu_1}{\mu_2} > \dfrac{\sigma_1^2}{\sigma_2^2} \end{cases} ,$$

$$r_2 > 0,\ R > 0$$

$$\begin{cases} -\Delta_1 R > r_2 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq 0 \geq \mu_2 \\ \mu_1 \neq \mu_2 \\ -\mu_2\sigma_1^2(1-2\sigma_2^2) > \mu_1\sigma_2^4 \\ \dfrac{2(\mu_1\sigma_2^2-\mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2-\sigma_2^2)^2} \leq \alpha \leq \dfrac{2\mu_2(\mu_2\sigma_1^2-\mu_1\sigma_2^2)}{\sigma_2^4} \end{cases} ,$$

and

$$
\begin{cases}
r_2 > 0,\ R > 0 \\
-\Delta_1 R > r_2 \\
\sigma_1^2 > \sigma_2^2 \\
0 > \mu_1 \geq \mu_2 \\
-\mu_2\sigma_1^2(1-2\sigma_2^2) > \mu_1\sigma_2^4 \\
\dfrac{2(\mu_1\sigma_2^2-\mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2-\sigma_2^2)^2} \leq \alpha \leq \dfrac{2\mu_2(\mu_2\sigma_1^2-\nu_1\sigma_2^2)}{\sigma_2^4}
\end{cases}
.
$$

With each of the parameter combinations in Group I, $F(0) > 0$ and $\lim_{z \uparrow \infty} F(z) < 0$. Hence there exists a positive solution z to the equation $F(z) = 0$. Call the first such root z_1, that is $z_1 > 0$ is such that $F(z_1) = 0$ and

$$
F(z) > 0 \ , \qquad \text{for all} \quad z \in [0, z_1) \ .
$$

Let G be the following function on $[0,\infty)$

$$
G(z) = (\rho-n)\ F(z) + F'(z) \tag{31}
$$

and upon substitution of (28) into (31), we find that G is everywhere negative for the parameter combinations in Group I. Therefore, $F'(z) < 0$ for all $z \in [0,z_1]$. Now suppose that there exists another solution to $F(z) = 0$. Call the next such root z_2, that is $z_2 > z_1$ is such that $F(z_2) = 0$ and

$$
F(z) \neq 0 \ , \qquad \text{for all} \quad z \in (z_1, z_2) \ .
$$

Thus it must be that

$$F(z) < 0 , \qquad \text{for all} \quad z \in (z_1, z_2) ,$$

and that $F'(z_2) \geq 0$. This however contradicts the negativity of $G(z_2)$, and we conclude that under the Group I cases there uniquely exists a positive solution z to (25). Similarly, with each of the parameter combinations in Group II we have that $F(0) < 0$, $\lim_{z \uparrow \infty} F(z) > 0$ and $G(z)$, as defined by (31), is everywhere positive on $[0,\infty)$. So again there exists a unique $z > 0$ satisfying (28).

Given our parameter restrictions (29) and (30) and the positive critical number z uniquely defined by (28), we now evaluate the function H for policy f_z. Expansion of (27) becomes

$$(32) \qquad H(x) = \frac{r_2}{\alpha}\left[\frac{(\rho-\eta)\, e^{-(\eta+\rho)z} + \beta(e^{-\eta z} - e^{-\rho z})}{\rho e^{-\rho z} - \eta e^{-\eta z} + \beta(e^{-\eta z} - e^{-\rho z})}\right]\left[\frac{\Delta_3\, e^{-\eta x} - \Delta_2\, e^{-\rho x}}{e^{-\eta z} - e^{-\rho z}}\right]$$

$$+ R\left[\frac{(\rho-\eta)}{\rho e^{\eta z} - \eta e^{\rho z} + \beta(e^{\rho z} - e^{\eta z})}\right]\left[\frac{\Delta_2\, e^{-\rho x} - \Delta_3\, e^{-\eta x}}{e^{-\eta z} - e^{-\rho z}}\right]$$

$$+ \left[R - \frac{r_2}{\alpha}\right]\left[\frac{\Delta_3\, e^{-\eta x-\rho z} - \Delta_2\, e^{-\rho x-\eta z}}{e^{-\eta z} - e^{-\rho z}}\right] + r_2 , \qquad \text{for} \quad 0 \leq x \leq z ,$$

and

$$(33) \qquad H(x) = \frac{r_2}{\alpha}\left[\frac{\rho e^{-\rho z} - \eta e^{-\eta z} - (\rho-\eta)\, e^{-(\eta+\rho)z}}{\rho e^{-\rho z} - \eta e^{-\eta z} + \beta(e^{-\eta z} - e^{-\rho z})}\right]\Delta_1\, e^{-\beta(x-z)}$$

$$+ R\left[\frac{(\rho-\eta)}{\rho e^{\eta z} - \eta e^{\rho z} + \beta(e^{\rho z} - e^{\eta z})}\right]\Delta_1\, e^{-\beta(x-z)} + r_2 ,$$

$$\text{for} \quad x \geq z .$$

Subsequently substitution of (28) into (32) and (33) results in $H(z) = 0$, and it is exactly our parameter situations (29) and (30) that guarantee the negativity of H on $[0,z)$ and the positivity of H on (z,∞). Thus the single critical number policy that uses control mode 1 whenever the state of the system is above level z and mode 2 when the state is below z, where z is the unique solution to the (28), is optimal for the following explicit parameter combinations:

$$[33] \quad \begin{cases} r_2 > 0,\ R > 0 \\ -\Delta_1 R > r_2 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq \mu_2 > 0 \\ \mu_1\sigma_2^2 - \mu_2\sigma_1^2 \leq 0 \end{cases} ,$$

$$[34] \quad \begin{cases} r_2 > 0,\ R > 0 \\ -\Delta_1 R > r_2 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq \mu_2 > 0 \\ \dfrac{\mu_1}{\mu_2} > \dfrac{\sigma_1^2}{\sigma_2^2} \\ \alpha \geq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2 - \sigma_2^2)^2} \end{cases} ,$$

$$[35]\quad \begin{cases} r_2 > 0,\ R > 0 \\ -\Delta_1 R > r_2 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq 0 \geq \mu_2 \\ \mu_1 \neq \mu_2 \\ \alpha \geq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2 - \sigma_2^2)^2} \end{cases} ,$$

$$[36]\quad \begin{cases} r_2 > 0,\ R > 0 \\ -\Delta_1 R > r_2 \\ \sigma_1^2 > \sigma_2^2 \\ 0 > \mu_1 \geq \mu_2 \\ \alpha \geq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2 - \sigma_2^2)^2} \end{cases} ,$$

$$[37]\quad \begin{cases} r_2 > 0,\ R < 0 \\ -\Delta_1 R > r_2 \\ \sigma_1^2 \leq \sigma_2^2 \end{cases} ,$$

$$[38]\quad \begin{cases} r_2 > 0,\ R < 0 \\ -\Delta_1 R > r_2 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq \mu_2 > 0 \\ \dfrac{\mu_1}{\mu_2} > \dfrac{\sigma_1^2}{\sigma_2^2} \\ \alpha \leq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2 - \sigma_2^2)^2} \end{cases}$$

$$[39]\quad \begin{cases} r_2 > 0,\ R < 0 \\ -\Delta_1 R > r_2 \\ \sigma_1^2 > \sigma_2^2 \\ \mu_1 \geq 0 \geq \mu_2 \\ \mu_1 \neq \mu_2 \\ \alpha \leq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2 - \sigma_2^2)^2} \end{cases},$$

and

$$[40]\quad \begin{cases} r_2 > 0,\ R < 0 \\ -\Delta_1 R > r_2 \\ \sigma_1^2 > \sigma_2^2 \\ 0 > \mu_1 \geq \mu_2 \\ \alpha \leq \dfrac{2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)\ (\mu_1-\mu_2)}{(\sigma_1^2 - \sigma_2^2)^2} \end{cases}$$

There then remains the case when operational cost r_2 is zero and operational cost r_1 is positive. The analysis is repititious of that discussed above for $r_1 = 0$ and $r_2 > 0$, and so will be omitted here. The results, however, are stated as follows. The single band policy of always using control mode 2 is optimal when the cost and diffusion parameters satisfy

$$(34)\quad \begin{cases} r_1 > 0,\ R \geq 0 \\ \Delta_2 \geq 0 \end{cases},$$

$$(35)\qquad \begin{cases} r_1 > 0,\ R \geq 0 \\ \Delta_2 < 0 \\ -\Delta_2 R \leq r_1\ , \end{cases}$$

$$(36)\qquad \begin{cases} r_1 > 0,\ R < 0 \\ \Delta_2 \leq 0\ , \end{cases}$$

or

$$(37)\qquad \begin{cases} r_1 > 0,\ R < 0 \\ \Delta_2 > 0 \\ -\Delta_2 R \leq r_1\ . \end{cases}$$

(The explicit enumeration of the parameter combinations that fall under (34) through (37) could likewise be carried out as was done for combinations [17] through [40].) For the remaining parameter situations

$$(38)\qquad \begin{cases} r_1 > 0,\ R \geq 0 \\ \Delta_2 < 0 \\ -\Delta_2 R > r_1\ , \end{cases}$$

and

$$(39)\qquad \begin{cases} r_1 > 0,\ R < 0 \\ \Delta_2 > 0 \\ -\Delta_2 R > r_1\ , \end{cases}$$

a single critical number policy is optimal. The optimal policy is to use control mode 2 whenever the state of the system is above level z

and mode 1 when the state is below z, where z is the unique positive solution to the transcendental equation

$$(40) \qquad \frac{r_1}{\alpha} \rho[\beta(\beta-\rho)e^{-\beta z} - \nu(\nu-\rho)e^{-\nu z}] + (R - \frac{r_1}{\alpha}) (\beta-\nu) (\rho-\beta) (\nu-\rho) e^{-(\nu+\beta)z} = 0 ,$$

and

$$\nu = \frac{\mu_1 - \sqrt{\mu_1^2 + 2\alpha\sigma_1^2}}{\sigma_1^2} .$$

5.3. Reflection and No Switching Costs

In this section we solve the two mode control problem with reflection at the boundary and zero switching costs. That is, we construct a stationary policy that is optimal and likewise compute the associated optimal return function. To emphasize the effect of linear holding costs in the reflecting barrier case, we will assume that the operational costs r_1 and r_2 are equal. Therefore as seen in Section 5.2 we may set r_1 and r_2 at zero.

We first look at the problem where the linear holding costs are incurred at a positive rate, and without loss of generality let $h = 1$. Again, we begin our analysis with the simple band policies f_1 (always use mode 1) and f_2 (always use mode 2). For policy f_1 and any initial state $x \in S$ we have

(41) $$V_{f_1}(x) = E\left[\int_0^{T_x} e^{-\alpha t} Z_1(t|x)dt + V_{f_1}(0)\, E[e^{-\alpha T_x}]\right] ,$$

where $Z_1(\cdot|x)$ is (unrestricted) Brownian Motion starting in x with drift μ_1 and variance parameter σ_1^2 and where $T_x = \inf\{t \geq 0 : Z_1(t|x) = 0\}$. (Recall that in the absence of switching costs the return function associated with any stationary policy is a function of state only.) We have already seen that

$$E[e^{-\alpha T_x}] = e^{-\beta x}$$

and that

$$E\left[\int_0^{\infty} e^{-\alpha t} Z_1(t|x)dt\right] = \frac{x}{\alpha} + \frac{\mu_1}{\alpha^2} .$$

Now since

$$E\left[\int_{T_x}^{\infty} e^{-\alpha t} Z_1(t|x)dt\right] = E\left[e^{-\alpha T_x} \int_0^{\infty} Z_1(t|0)dt\right] ,$$

it must be that

$$E\left[\int_0^{T_x} e^{-\alpha t} Z_1(t|x)dt\right] = \frac{x}{\alpha} + \frac{\mu_1}{\alpha^2} - \frac{\mu_1}{\alpha^2} e^{-\beta x} , \qquad \text{for all } x \in S .$$

Hence

(42) $$V_{f_1}(x) = \frac{x}{\alpha} + \frac{\mu_1}{\alpha^2} + \left[V_{f_1}(0) - \frac{\mu_1}{\alpha^2}\right] e^{-\beta x} , \qquad \text{for all } x \in S ,$$

and similarly

(43) $$V_{f_2}(x) = \frac{x}{\alpha} + \frac{\mu_2}{\alpha^2} + \left[V_{f_2}(0) - \frac{\mu_2}{\alpha^2}\right] e^{-\rho x} , \qquad \text{for all } x \in S .$$

We can determine the value of $V_{f_1}(0)$ and $V_{f_2}(0)$ by applying boundary condition (36) from Theorem 4 in Chapter 3 to return functions V_{f_1} and V_{f_2}, respectively. The boundary conditions here are

(44) $$V'_{f_1}(0) = V'_{f_2}(0) = 0 \quad ,$$

which result in

(45) $$V_{f_1}(0) = \frac{1}{\alpha\beta} + \frac{\mu_1}{\alpha^2} \quad ,$$

(46) $$V_{f_1}(x) = \frac{x}{\alpha} + \frac{\mu_1}{\alpha^2} + \frac{1}{\alpha\beta} e^{-\beta x} \quad , \qquad \text{for all } x \in S \ ,$$

(47) $$V_{f_2}(0) = \frac{1}{\alpha\rho} + \frac{\mu_2}{\alpha^2} \quad ,$$

and

(48) $$V_{f_2}(x) = \frac{x}{\alpha} + \frac{\mu_2}{\alpha^2} + \frac{1}{\alpha\rho} e^{-\rho x} \quad , \qquad \text{for all } x \in S \ .$$

This leads to the following relationships on S

(49) $$D_2 V_{f_1}(x) - \alpha V_{f_1}(x) + x = \frac{(\mu_2-\mu_1)}{\alpha} + \frac{1}{\alpha\beta} \Delta_1 e^{-\beta x} \quad ,$$

and

(50) $$D_1 V_{f_2}(x) - \alpha V_{f_2}(x) + x = \frac{(\mu_1-\mu_2)}{\alpha} + \frac{1}{\alpha\rho} \Delta_2 e^{-\rho x} \quad .$$

Therefore the (necessary and sufficient) optimality conditions are satisfied for policy f_1 if and only if

(51) $$\begin{cases} \mu_1 = \mu_2 \\ \Delta_1 \geq 0 \end{cases} ,$$

and for policy f_2 if and only if

(52) $$\Delta_2 \geq (\mu_2 - \mu_1)\rho \quad .$$

Further inspection of (51) and (52) shows that it is optimal to always use mode 1 if we have the parameter combination

[41] $$\begin{matrix} \mu_1 = \mu_2 \\ \sigma_1^2 \leq \sigma_2^2 \end{matrix} \quad ,$$

while it is optimal to always use mode 2 if we have

[42] $$\sigma_1^2 \geq \sigma_2^2 \quad .$$

There then remains the situation where $\mu_1 > \mu_2$ and $\sigma_1^2 < \sigma_2^2$. Under such circumstances we think of mode 1 as the "tortoise" and mode 2 as the "hare", and we call our control problem a <u>tortoise-hare problem</u>. This tortoise-hare arrangement suggests that we next investigate the single critical number policy f_z given by

$$f_z(x,1) = f_z(x,2) = \begin{cases} 1, & \text{if } x \in [0,z] \\ 2, & \text{if } x \in [z,\infty). \end{cases}$$

Using the Markov arguments of Section 5.2 that led to (20) and those above that led to (42) and (43), we see that

$$(53)\qquad V_{f_z}(x) = \frac{x}{\alpha} + [1 - \varphi(x) - \psi(x)]\,\frac{\mu_1}{\alpha^2} - \varphi(x)\,\frac{z}{\alpha} + \varphi(x)\,V_{f_z}(z) + \psi(x)\,V_{f_z}(0),$$

$$\text{for } x \in [0,z]$$

and

$$(54)\qquad V_{f_z}(x) = \frac{x}{\alpha} + [1-\delta(x)]\,\frac{\mu_2}{\alpha^{\delta}} - \delta(x)\,\frac{z}{\alpha} + \delta(x)\,V_{f_z}(z)\ , \quad \text{for } x \in [z,\infty)\ ,$$

where

$$\varphi(x) = \frac{e^{-\nu x} - e^{-\beta x}}{e^{-\nu z} - e^{-\beta z}}\ , \qquad \text{for } x \in [0,z]\ ,$$

$$\psi(x) = \frac{e^{-\nu z-\beta x} - e^{-\beta z-\nu x}}{e^{-\nu z} - e^{-\beta z}}\ , \qquad \text{for } x \in [0,z]\ ,$$

and

$$\delta(x) = e^{-\rho x}\ , \qquad \text{for } x \in [z,\infty)\ .$$

From Theorem 4, Chapter 3 we find that $V'_{f_z}(x)$ exists everywhere on S and is zero at $x = 0$. Thus we have the following two equations in the two unknowns $V_{f_z}(0)$ and $V_{f_z}(z)$

$$(55)\qquad \frac{1}{\alpha} + \left[V_{f_z}(z) - \frac{z}{\alpha} - \frac{\mu_1}{\alpha^2}\right]\varphi'(0) + \left[V_{f_z}(0) - \frac{\mu_1}{\alpha^2}\right]\psi'(0) = 0$$

and

$$(56)\qquad \left[V_{f_z}(z) - \frac{z}{\alpha} - \frac{\mu_1}{\alpha^2}\right]\varphi'(z) + \left[V_{f_z}(0) - \frac{\mu_1}{\alpha^2}\right]\psi'(z)$$

$$- \left[V_{f_z}(z) - \frac{z}{\alpha} - \frac{\mu_2}{\alpha^2}\right]\delta'(z) = 0\ ,$$

which we solve for

$$(57)\quad V_{f_z}(0) = \frac{\mu_1}{\alpha^2}\,[\nu\beta(e^{(\nu-\beta)z} + e^{(\beta-\nu)z}) + \rho e^{-\rho z}(e^{-\nu z} - e^{-\beta z})(\nu e^{\nu z} - \beta e^{\beta z})]^{-1}$$

$$\cdot\ \{\nu\beta(e^{(\nu-\beta)z} + e^{(\beta-\nu)z}) + \rho e^{-\rho z}(e^{-\nu z} - e^{-\beta z})(\nu e^{\nu z} - \beta e^{\beta z})$$

$$+ (\beta-\nu)\ \rho e^{-\rho z}(e^{\beta z} - e^{\nu z})\}$$

$$- \frac{\mu_2}{\alpha^2}\,[\nu\beta(e^{(\nu-\beta)z} + e^{(\beta-\nu)z}) + \rho e^{-\rho z}(e^{-\nu z} - e^{-\beta z})(\nu e^{\nu z}-\beta e^{\beta z})]^{-1}$$

$$\cdot\ (\beta-\nu)\ \rho e^{-\rho z}(e^{\beta z} - e^{\nu z})$$

$$- \frac{1}{\alpha}\,[\nu\beta(e^{(\nu-\beta)z} + e^{(\beta-\nu)z})\ (\nu e^{-\beta z} - \beta e^{-\nu z})$$

$$+ \rho e^{-\rho z}(e^{-\nu z} - e^{-\beta z})(\nu e^{\nu z} - \beta e^{\beta z})(\nu e^{-\beta z} - \beta e^{-\nu z})]^{-1}$$

$$\cdot\ \{\nu\beta(e^{-\nu z} - e^{-\beta z})\ (e^{(\nu-\beta)z} + e^{(\beta-\nu)z}) - (\beta-\nu)(\nu-\beta)(e^{-\nu z}-e^{-\beta z})$$

$$+ \rho e^{-\rho z}(e^{-\nu z} - e^{-\beta z})^2\ (\nu e^{\nu z}-\beta e^{\beta z})\}\ ,$$

and

$$(58)\quad V_{f_z}(z) = \frac{\mu_1}{\alpha^2}\,[\nu\beta(e^{(\nu-\beta)z} + e^{(\beta-\nu)z}) + \rho e^{-\rho z}(e^{-\nu z} - e^{-\beta z})\ (\nu e^{\nu z} - \beta e^{\beta z})]^{-1}$$

$$\cdot\ \nu\beta(e^{(\nu-\beta)z} + e^{(\beta-\nu)z}) + \frac{z}{\alpha}$$

$$+ \frac{\mu_2}{\alpha^2}\,[\nu\beta(e^{(\nu-\beta)z} + e^{(\beta-\nu)z}) + \rho e^{-\rho z}(e^{-\nu z} - e^{-\beta z})(\nu e^{\nu z}-\beta e^{\beta z})]^{-1}$$

$$\cdot\ \rho e^{-\rho z}(e^{-\nu z} - e^{-\beta z})\ (\nu e^{\nu z} - \beta e^{\beta z})$$

$$+ \frac{1}{\alpha}\,[\nu\beta(e^{(\nu-\beta)z} + e^{(\beta-\nu)z}) + \rho e^{-\rho z}(e^{-\nu z} - e^{-\beta z})(\nu e^{\nu z}-\beta e^{\beta z})]^{-1}$$

$$\cdot\ (\nu-\beta)\ (e^{-\nu z} - e^{-\beta z})\ .$$

The explicit solution of V_{f_z} on S, then, is as follows

$$(59) \quad V_{f_z}(x) = \frac{x}{\alpha} + \frac{\mu_1}{\alpha^2}[e^{-\nu z} - e^{-\beta z}]^{-1}$$

$$\cdot \{(e^{-\nu z} - e^{-\beta z}) + (e^{-\nu x} - e^{-\beta x})[\bar{A}(z)-1]$$

$$+ (e^{-\nu z-\beta x} - e^{-\beta z-\nu x})[\bar{D}(z)-1]\}$$

$$+ \frac{\mu_2}{\alpha^2}[e^{-\nu z} - e^{-\beta z}]^{-1} \cdot$$

$$\cdot \{(e^{-\nu x} - e^{-\beta x})\,\bar{B}(z) - (e^{-\nu z-\beta x} - e^{-\beta z-\nu x})\,\bar{E}(z)\}$$

$$+ \frac{1}{\alpha}[e^{-\nu z} - e^{-\beta z}]^{-1}$$

$$\cdot \{(e^{-\nu x} - e^{-\beta x})\,\bar{C}(z) - (e^{-\nu z-\beta x} - e^{-\beta z-\nu x})\,\bar{F}(z)\} \quad , \quad \text{for} \quad 0 \leq x \leq z$$

and

$$(60) \quad V_{f_z}(x) = \frac{x}{\alpha} + \frac{\mu_1}{\alpha^2} e^{-\rho x}\bar{A}(z) + \frac{\mu_2}{\alpha^2}[1 - e^{-\rho x} + e^{-\rho x}\bar{B}(z)]$$

$$+ \frac{1}{\alpha} e^{-\rho x}[\bar{C}(z)-1] \quad , \qquad \text{for} \quad x \geq z \; ,$$

where

$$\bar{A}(z) = [\nu\beta(e^{(\nu-\beta)z} + e^{(\beta-\nu)z}) + \rho e^{-\rho z}(e^{-\nu z} - e^{-\beta t})\,(\nu e^{\nu z} - \beta e^{\beta z})]^{-1}$$

$$\cdot \nu\beta(e^{(\nu-\beta)z} + e^{(\beta-\nu)z}) \; ,$$

$$\bar{B}(z) = [\nu\beta(e^{(\nu-\beta)z} + e^{(\beta-\nu)z}) + \rho e^{-\rho z}(e^{-\nu z} - e^{-\beta z})\,(\nu e^{\nu z} - \beta e^{\beta z})]^{-1}$$

$$\cdot \rho e^{-\rho z}(e^{-\nu z} - e^{-\beta z})\,(\nu e^{\nu z} - \beta e^{\beta z}) \quad ,$$

$$\bar{C}(z) = [\nu\beta(e^{(\nu-\beta)z} + e^{(\beta-\nu)z}) + \rho e^{-\rho z}(e^{-\nu z} - e^{-\beta z})\,(\nu e^{\nu z} - \beta e^{\beta z})]^{-1}$$

$$\cdot\,(\nu-\beta)\,(e^{-\nu z} - e^{-\beta z}) \quad ,$$

$$\bar{D}(z) = [\nu\beta(e^{(\nu-\beta)z} + e^{(\beta-\nu)z}) + \rho e^{-\rho z}(e^{-\nu z} - e^{-\beta z})\,(\nu e^{\nu z} - \beta e^{\beta z})]^{-1}$$

$$\cdot\,[\nu\beta(e^{(\nu-\beta)z} + e^{(\beta-\nu)z}) + \rho e^{-\rho z}(e^{-\nu z} - e^{-\beta z})\,(\nu e^{\nu z} - \beta e^{\beta z})$$

$$+ (\beta-\nu)\,\rho e^{-\rho z}(e^{\beta z} - e^{\nu z})] \quad ,$$

$$\bar{E}(z) = [\nu\beta(e^{(\nu-\beta)z} + e^{(\beta-\nu)z} + \rho e^{-\rho z}(e^{-\nu z} - e^{-\beta z})\,(\nu e^{\nu z} - \beta e^{\beta z})]^{-1}$$

$$\cdot\,(\beta-\nu)\,\rho e^{-\rho z}(e^{\beta z} - e^{\nu z}) \quad ,$$

and

$$\bar{F}(z) = [\nu e^{-\beta z} - \beta e^{-\nu z}]^{-1}\,(\beta-\nu) \quad .$$

As in Section 5.2 the optimality conditions now lead us to choose our critical number so that $V''_{f_z}(z-) = V''_{f_z}(z+)$. Therefore we wish z to satisfy the following equation

$$(61)\quad \left[\frac{\mu_1}{\alpha^2} - \frac{\mu_2}{\alpha^2}\right]\{[e^{-\nu z} - e^{-\beta z}]^{-1}\,(\beta e^{-\beta z} - \nu e^{-\nu z})\,[\bar{A}(z)-1] - \rho^2\,e^{-\rho z}\,\bar{A}(z)$$

$$+ [e^{-\nu z} - e^{-\beta z}]^{-1}\,(\nu-\beta)\,e^{-(\nu+\beta)z}\,[\bar{D}(z)-1]\}$$

$$+ \frac{1}{\alpha}\{[e^{-\nu z} - e^{-\beta z}]^{-1}\,(\beta e^{-\beta z} - \nu e^{-\nu z})\,\bar{C}(z) - \rho^2\,e^{-\rho z}\,[\bar{C}(z)-1]$$

$$+ [e^{-\nu z} - e^{-\beta z}]^{-1}\,(\beta-\nu)\,e^{-(\nu+\beta)z}\,\bar{F}(z)\} = 0 \quad .$$

Call the left hand side of (61) $F(z)$ where $z \in [0,\infty)$, and let G be the function on $[0,\infty)$

$$G(z) = (\beta-\nu)\ F(z) + F\ (z)\quad .$$

With much pencil pushing it can be shown that $F(0) > 0$, $\lim_{z \uparrow \infty} F(z) = -\infty$, and $G(z) < 0$ for all $z \in [0,\infty)$ when $\mu_1 > \mu_2$, $\sigma_1^2 < \sigma_2^2$ and $\alpha > \mu_1 - \mu_2$. Similarly $F(0) < 0$, $\lim_{z \uparrow \infty} F(z) = +\infty$, and $G(z) > 0$ for all $z \in [0,\infty)$ when $\mu_1 > \mu_2$, $\sigma_1^2 < \sigma_2^2$ and $0 < \alpha < \mu_1 - \mu_2$. Thus as argued in Section 5.2, equation (61) $(F(z) = 0)$ has a unique positive solution z^*. Taking this z^* we substitute back into (59), (60) and the related expressions $\bar{A}(z^*)$ through $\bar{F}(z^*)$. Let H be the following function on S

$$(62)\qquad H(x) = D_2\, V_{f_{z^*}}(x) - D_1\, V_{f_{z^*}}(x)\ .$$

When expanded definition (62) becomes

$$
\begin{aligned}
(63)\quad H(x) = {} & \frac{(\mu_2-\mu_1)}{\alpha^2}\,[e^{-\nu z^*} - e^{-\beta z^*}]^{-1}\,\bar{B}(z^*)\,\{\Delta_4\, e^{-\nu x} - \Delta_1\, e^{-\beta x}\} \\
& + \frac{(\mu_2-\mu_1)}{\alpha^2}\,[e^{-\nu z^*} - e^{-\beta z^*}]^{-1}\,\bar{E}(z^*)\,\{\Delta_4\, e^{-\beta z^* - \nu x} - \Delta_1 e^{-\nu z^* - \beta x}\} \\
& + \frac{1}{\alpha}\,[e^{-\nu z^*} - e^{-\beta z^*}]^{-1}\,\bar{C}(z^*)\,\{\Delta_4\, e^{-\nu x} - \Delta_1\, e^{-\beta x}\} \\
& + \frac{1}{\alpha}\,[e^{-\nu z^*} - e^{-\beta z^*}]^{-1}\,\bar{F}(z^*)\,\{\Delta_4\, e^{-\beta z^* - \nu x} - \Delta_4\, e^{-\nu z^* - \beta x}\} \\
& + \frac{(\mu_2-\mu_1)}{\alpha}\ , \qquad\qquad \text{for } 0 \le x \le z^*
\end{aligned}
$$

and

$$(64) \qquad H(x) = \frac{(\mu_2-\mu_1)}{\alpha^2} \bar{A}(z^*) \Delta_2 e^{-\rho x} + \frac{1}{\alpha} [1 - \bar{C}(z^*)] \Delta_2 e^{-\rho x} + \frac{(\mu_2-\mu_1)}{\alpha^2} ,$$

$$\text{for } x \geq z^* ,$$

where we have added to the previous parameter functions Δ_1, Δ_2 and Δ_3, the function

$$\Delta_4 = \frac{1}{2} \sigma_2^2 \nu^2 - \mu_2 \nu - \alpha \ .$$

Inspection of (63) and (64) shows that when $\mu_1 > \mu_2$ and $\sigma_1^2 < \sigma_2^2$, we have that $H(x) > 0$ for $x \in [0,z^*)$, $H(z^*) = 0$ and $H(x) < 0$ for $x \in (z^*,\infty)$. Thus we conclude that $D_2 V_{f_{z^*}}(x) - \alpha V_{f_{z^*}}(x) + x \geq 0$ on $[0,z^*]$ and $D_1 V_{f_{z^*}}(x) - \alpha V_{f_{z^*}}(x) + x \geq 0$ on $[z^*,\infty)$; and the single critical number policy that uses control mode 2 whenever the state of the system is above level z^*, the unique solution to (61), and uses control mode 1 when the state is below z^* is optimal for our tortoise-hare problem.

If instead the linear holding costs are incurred at a negative rate and we re-scale our system so that $h = -1$, we find as expected that the policy of always using mode 1 is optimal if $\sigma_1^2 \geq \sigma_2^2$, while the policy of always using mode 2 is optimal only if $\mu_1 = \mu_2$ and $\sigma_1^2 \leq \sigma_2^2$. Finally for the tortoise-hare situation $\mu_1 > \mu_2$ and $\sigma_1^2 < \sigma_2^2$, the optimal policy is a single critical number policy where mode 1 is

used whenever the state of the system is above the critical level and mode 2 is used when the state is below.

5.4. Reflection and Switching Costs

We would now like to investigate the effect of switching costs on our solutions in the reflecting barrier control problem. To keep our calculations somewhat manageable we will specifically treat the tortoise-hare situation $(\mu_1 > \mu_2$ and $\sigma_1^2 < \sigma_2^2)$ and assume equal operational costs (WLOG $r_1 = r_2 = 0$), positive holding costs (WLOG $h = 1$), and a symmetric switching cost $K_{12} = K_{21} + K > 0$. Note that with the addition of switching costs, the return function corresponding to any stationary policy is a function of both state and mode.

As before, we look first at the simple band policies f_1 (always use mode 1) and f_2 (alwyas use mode 2). Referring to (46) and (48) we have that

$$(65) \qquad V_{f_1}(x,1) = \frac{x}{\alpha} + \frac{\mu_1}{\alpha^2} + \frac{1}{\alpha\beta} e^{-\beta x}, \qquad x \in S,$$

$$(66) \qquad V_{f_1}(x,2) = K + \frac{x}{\alpha} + \frac{\mu_1}{\alpha^2} + \frac{1}{\alpha\beta} e^{-\beta x}, \qquad x \in S,$$

$$(67) \qquad V_{f_2}(x,1) = K + \frac{x}{\alpha} + \frac{\mu_2}{\alpha^2} + \frac{1}{\alpha\rho} e^{-\rho x}, \qquad x \in S,,$$

and

$$(68) \qquad V_{f_2}(x,2) = \frac{x}{\alpha} + \frac{\mu_2}{\alpha^2} + \frac{1}{\alpha\rho} e^{-\rho x}, \qquad x \in S.$$

Optimality condition (13) of Theorem 1 in Chapter 4 is immediately satisfied by both f_1 and f_2. If

$$(69) \quad D_2 V_{f_1}(x,2) - \alpha V_{f_1}(x,2) + x = \frac{(\mu_2-\mu_1)}{\alpha} + \Delta_1 \frac{1}{\alpha\beta} e^{-\beta x} - \alpha K \geq 0 \quad ,$$

for $x \in S$,

then condition (14) of the same theorem will hold for policy f_1, and if

$$(70) \quad D_1 V_{f_2}(x,1) - \alpha V_{f_2}(x,1) + x = \frac{(\mu_1-\mu_2)}{\alpha} + \Delta_2 \frac{1}{\alpha\rho} e^{-\rho x} - \alpha K \geq 0 \quad ,$$

for $x \in S$,

it will hold for policy f_2. Since $\sigma_1^2 < \sigma_2^2$ implies that $\Delta_1 > 0$ and $\Delta_2 < 0$, (69) and (70) will be satisfied if

$$(71) \qquad \mu_2-\mu_1 \geq \alpha^2 K$$

and

$$(72) \qquad \mu_1-\mu_2 + \Delta_2 \frac{1}{\rho} \geq \alpha^2 K \quad ,$$

respectively. That switching cost K is strictly positive contradicts (71) for $\mu_1 > \mu_2$, and the fact that

$$\mu_1-\mu_2 + \Delta_2 \frac{1}{\rho} = -\mu_2 + \frac{1}{2}\sigma_1^2\rho - \frac{\alpha}{\rho}$$

$$= \frac{(\sigma_1^2-\sigma_2^2)\left(\mu_2 + \sqrt{\mu_2^2 + 2\alpha\sigma_2^2}\right)^2}{2\sigma_2^2} < 0$$

likewise contracits (72) for $\sigma_1^2 < \sigma_2^2$. Therefore we have shown that with a tortoise and hare alignment of control modes, it is not optimal to strictly use either mode.

The results of Section 5.3 suggest that we now restrict attention to policies f_* of the following form

$$f_*(x,1) = \begin{cases} 1, & \text{if } x \in [0,Z) \\ 2, & \text{if } x \in [Z,\infty) \end{cases}$$

and

$$f_*(x,2) = \begin{cases} 1, & \text{if } x \in [0,z] \\ 2, & \text{if } x \in (z,\infty), \end{cases}$$

where $0 \leq z < Z \leq \infty$. Such a policy is called a two critical numbers policy and is characterized by the values of the two switching levels z and Z.

Suppose that $z = 0$ and $Z = \infty$. Then policy f_* will continue with the initial control mode forever and

$$(73) \qquad V_{f_*}(x,1) = \frac{x}{\alpha} + \frac{\mu_1}{\alpha^2} + \frac{1}{\alpha\beta} e^{-\beta x}, \qquad \text{for all } x \in S$$

and

$$(74) \qquad V_{f_*}(x,2) = \frac{x}{\alpha} + \frac{\mu_2}{\alpha^2} + \frac{1}{\alpha\rho} e^{-\rho x}, \qquad \text{for all } x \in S.$$

Since

$$D_1 V_{f_*}(x,1) - \alpha V_{f_*}(x,1) + x = [\tfrac{1}{2}\sigma_1^2 \beta^2 - \mu_1\beta - \alpha] \frac{1}{\alpha\beta} e^{-\beta x} = 0 , \quad x \in S$$

and

$$D_2 V_{f_*}(x,2) - \alpha V_{f_*}(x,2) + x = [\tfrac{1}{2}\sigma_2^2 \rho^2 - \mu_2\rho - \alpha] \frac{1}{\alpha\rho} e^{-\rho x} = 0 , \quad x \in S,$$

optimality condition (14) of Chapter 4 holds everywhere for all possible parameters. However, optimality condition (13) of Chapter 4 will hold if and only if both of the following are satisfied,

$$(75) \qquad V_{f_*}(x,1) - V_{f_*}(x,2) = \frac{(\mu_1-\mu_2)}{\alpha^2} + \frac{\rho e^{\rho x} - \beta e^{\beta x}}{\alpha\beta\rho\, e^{(\rho+\beta)x}} \leq K , \quad \text{for all } x \in S$$

and

$$(76) \qquad V_{f_*}(x,2) - V_{f_*}(x,1) = \frac{(\mu_2-\mu_1)}{\alpha^2} + \frac{\beta e^{\beta x} - \rho e^{\rho x}}{\alpha\beta\rho\, e^{(\rho+\beta)x}} \leq K , \quad \text{for all } x \in S.$$

We evaluate (75) and (76) separately for the cases $\rho \leq \beta$ and $\rho > \beta$ and find inequalities (75) and (76) to hold everywhere if

$$\begin{cases} \rho \leq \beta \\ \mu_1-\mu_2 \leq \alpha^2 K \\ \mu_2-\mu_1 - \dfrac{\alpha(\rho-\beta)}{\rho\beta} \leq \alpha^2 K \end{cases}$$

or

$$\begin{cases} \rho > \beta \\ \mu_1-\mu_2 + \dfrac{\alpha(\rho-\beta)}{\rho\beta} \leq \alpha^2 K \\ \mu_2-\mu_1 \leq \alpha^2 K \end{cases}$$

hold. Thus we conclude that the policy of never switching control mode will be optimal if our cost and diffusion parameters fall into one of the following situations

$$[43] \quad \begin{cases} \mu_1 > \mu_2 \\ \sigma_1^2 < \sigma_2^2 \\ \rho \leq \beta \\ \mu_1-\mu_2 \leq \alpha^2 K \\ \mu_2-\mu_1 - \dfrac{\alpha(\rho-\beta)}{\rho\beta} \leq \alpha^2 K \end{cases}$$

and

$$[44] \quad \begin{cases} \mu_1 > \mu_2 \\ \sigma_1^2 < \sigma_2^2 \\ \rho > \beta \\ \mu_1-\mu_2 + \dfrac{\alpha(\rho-\beta)}{\rho\beta} \leq \alpha^2 K \end{cases} .$$

Suppose now that $z = 0$ and $0 < Z < \infty$. Previous arguments (see Sections 5.2 and 5.3) then lead to

$$
(76)\quad V_{f_*}(x,1) = \begin{cases} [1-\varphi(x)-\psi(x)]\,\dfrac{\mu_1}{\alpha^2} + \varphi(x)\,\dfrac{\mu_2}{\alpha^2} + \dfrac{x}{\alpha} + \varphi(x)K + \varphi(x)\,\dfrac{1}{\alpha\rho}\,e^{-\rho z} \\ \qquad\qquad + \psi(x)\,V_{f_*}(0,1), & x \in [0,Z) \\ \\ K + \dfrac{x}{\alpha} + \dfrac{\mu_2}{\alpha^2} + \dfrac{1}{\alpha\rho}\,e^{-\rho x}, & x \in [Z,\infty) \end{cases}
$$

and

$$
(77)\quad V_{f_*}(x,z) = \frac{x}{\alpha} + \frac{\mu_2}{\alpha^2} + \frac{1}{\alpha\rho}\,e^{-\rho x}, \qquad x \in S,
$$

where φ and ψ are the following functions on $[0,Z]$

$$
\varphi(x) = \frac{e^{-\nu x} - e^{-\beta x}}{e^{-\nu Z} - e^{-\beta Z}},
$$

and

$$
\psi(x) = \frac{e^{-\nu Z-\beta x} - e^{-\beta Z-\nu x}}{e^{-\nu Z} - e^{-\beta Z}}.
$$

Our characterization of return functions (Theorem 4, Chapter 3) states that $V'_{f_*}(0,1) = 0$, and so we have

$$
\begin{aligned}
(78)\quad V_{f_*}(0,1) = {} & \frac{\mu_1}{\alpha^2}[\psi'(0)]^{-1}[\varphi'(0)+\psi'(0)] - \frac{\mu_2}{\alpha^2}[\psi'(0)]^{-1}\varphi'(0) \\
& - \frac{1}{\alpha}[\psi'(0)]^{-1} - K[\psi'(0)]^{-1}\varphi'(0) - \frac{1}{\alpha}[\psi'(0)\rho]^{-1}\varphi'(0)e^{-\rho Z}
\end{aligned}
$$

which when entered into (76) completes our derivation of the return function in terms of critical number Z. (We have refrained here from

writing out in length all of the exponential-type expressions involved.)

If such a policy f_* is to be optimal, then the optimality conditions include the requirement that $D_1 V_{f_*}(x,1) - \alpha V_{f_*}(x,1) + x \geq 0$ for all $x \in [Z,\infty)$. Thus we need that

$$\frac{(\mu_1-\mu_2)}{\alpha} + \Delta_2 \frac{1}{\alpha\rho} e^{-\rho x} \geq \alpha K \quad , \qquad \text{for all } x \in [Z,\infty) \; ,$$

which can be the case only if $\mu_1-\mu_2 > \alpha^2 K$. In order for $V_{f_*}(x,1) \leq K + V_{f_*}(x,2)$ and $V_{f_*}(x,2) \leq K + V_{f_*}(x,1)$ on S, we need satisfy

$$(79) \quad \left| \frac{(\mu_1-\mu_2)}{\alpha^2}[1-\varphi(x)] - \frac{\mu_1}{\alpha^2}\psi(x) + K\varphi(x) + \frac{1}{\alpha\rho}[\varphi(x)e^{-\rho Z} - e^{-\rho x}] + V_{f_*}(0,1)\,\psi(x) \right| \leq K \; , \qquad \text{for all } x \in [0,Z] \; .$$

Inspection of (79) shows that it will hold true only if $\rho \leq \beta$, and so we anticipate that a two critical numbers policy of the form $0 = z < Z < \infty$ will be optimal only under the following parameter situation

$$[45] \quad \begin{cases} \mu_1 > \mu_2 \\ \sigma_1^2 < \sigma_2^2 \\ \rho \leq \beta \\ \mu_1-\mu_2 > \alpha^2 K \; . \end{cases}$$

Restricting attention to [45] let F be the function on S defined by

$$F(Z) = \mu_1-\mu_2 + (\tfrac{1}{2}\,\sigma_1^2\rho - \mu_1 - \tfrac{\alpha}{\rho})\, e^{-\rho Z}\,, \qquad Z \in [0,\infty)\,.$$

We have already seen that $F(0) < 0$ and $\lim\limits_{Z \uparrow \infty} F(Z) = \mu_1-\mu_2 > \alpha^2K$, and therefore there uniquely exists $Z^* > 0$ such that $F(Z^*) = \alpha^2K$. We now compute explicitly $V_{f_*}(0,1)$ and return function V_{f_*} on $S\times A$ using this critical number Z^*, and find that we have strict inequality in (79) on $[0,Z^*)$. The optimality conditions remaining to be verified are that $D_1 V_{f_*}(x,1) - \alpha V_{f_*}(x,1) + x = 0$ for all $x \in [0,Z^*]$ and that $D_2 V_{f_*}(x,2) - \alpha V_{f_*}(x,2) + x = 0$ for all $x \in S$, which are easily true by Theorem 4 of Chapter 3. Hence for parameter combination [45] a two critical numbers policy is optimal where $z = 0$ and $Z \in (0,\infty)$ is the unique solution to a transcendental equation.

We are left with two arrangements of parameters to account for, namely

$$[46] \quad \begin{cases} \mu_1 > \mu_2 \\ \sigma_1^2 < \sigma_2^2 \\ \rho \le \beta \\ \mu_1-\mu_2 \le \alpha^2K \\ \mu_2-\mu_1 - \dfrac{\alpha(\rho-\beta)}{\rho\beta} > \alpha^2K \end{cases}$$

and

$$[47] \quad \begin{cases} \mu_1 > \mu_2 \\ \sigma_1^2 < \sigma_2^2 \\ \rho > \beta \\ \mu_1-\mu_2 + \dfrac{\alpha(\rho-\beta)}{\rho\beta} > \alpha^2K \end{cases}$$

If $0 < z < Z < \infty$, then our return function in general form is given by

$$
(80) \quad V_{f_*}(x,1) = \begin{cases} [1-\varphi(x)-\psi(x)]\,\dfrac{\mu_1}{\alpha^2} - \varphi(x)\,\dfrac{Z}{\alpha} + \dfrac{x}{\alpha} + \varphi(x)[K+V_{f_*}(Z,2)] \\ \qquad\qquad + \psi(x)\,V_{f_*}(0,1)\,, \quad x \in [0,Z) \\ K + [1-\delta(x)]\,\dfrac{\mu_2}{\alpha^2} - \delta(x)\,\dfrac{z}{\alpha} + \dfrac{x}{\alpha} + \delta(x)[K+V_{f_*}(z,1)]\,, \\ \qquad\qquad x \in [Z,\infty)\,, \end{cases}
$$

and

$$
(81) \quad V_{f_*}(x,2) = \begin{cases} K+[1-\varphi(x)-\psi(x)]\,\dfrac{\mu_1}{\alpha^2} - \varphi(x)\,\dfrac{Z}{\alpha} + \dfrac{x}{\alpha} + \varphi(x)[K+V_{f_*}(Z,2)] \\ \qquad\qquad + \psi(x)\,V_{f_*}(0,1), \quad x \in [0,z] \\ [1-\delta(x)]\,\dfrac{\mu_2}{\alpha^2} - \delta(x)\,\dfrac{z}{\alpha} + \dfrac{x}{\alpha} + \delta(x)[K+V_{f_*}(z,1)], \quad x \in (z,\infty)\,, \end{cases}
$$

where additionally δ is defined on $[z,\infty)$ as

$$
\delta(x) = e^{-\rho(x-z)}\,, \qquad \text{for all } x \geq z\,.
$$

Imposing our boundary condition, $V'_{f_*}(0,1) = 0$, we then have the following three equations

$$
(82) \quad V_{f_*}(z,1) = [1-\varphi(z)-\psi(z)]\,\frac{\mu_1}{\alpha^2} - \varphi(z)\,\frac{Z}{\alpha} + \frac{z}{\alpha} + \varphi(z)[K+V_{f_*}(Z,2)]+\psi(z)V_{f_*}(0,1),
$$

$$
(83) \quad V_{f_*}(Z,2) = [1-\delta(Z)]\,\frac{\mu_2}{\alpha^2} - \delta(Z)\,\frac{z}{\alpha} + \frac{Z}{\alpha} + \delta(Z)[K+V_{f_*}(z,1)]\,,
$$

and

$$(84)\qquad 0 = \frac{1}{\alpha} + \varphi'(0)\left[K + V_{f_*}(Z,2) - \frac{Z}{\alpha} - \frac{\mu_1}{\alpha^2}\right] + \psi'(0)\left[V_{f_*}(0,1) - \frac{\mu_1}{\alpha^2}\right]$$

which can be solved for the unknowns $V_{f_*}(0,1)$, $V_{f_*}(z,1)$, and $V_{f_*}(Z,2)$. Substitutions of these three values into (80) and (81), then, leaves return function V_{f_*} in terms of unspecified z and Z.

If we now set out to minimize $V_{f_*}(x,1)$ and $V_{f_*}(x,z)$ over all possible (z,Z) pairs and for all $x \in S$, we find that when [46] and [47] hold our task is equivalent to minimizing the following two expressions with respect to z and Z,

$$\begin{aligned}(85)\quad \bar{G}(z,Z) = \Big\{ & [e^{-\nu Z} - e^{-\beta Z} - e^{-\rho(Z-z)-\nu z} + e^{-\rho(Z-z)-\beta z}]\,[1 - \nu(e^{-\nu Z} - e^{-\beta Z})] \\ & + (\beta-\nu)\,[e^{-\nu Z} - e^{-\beta Z} - e^{-\rho(Z-z)-\nu z} + e^{-\rho(Z-z)-\beta z}] \\ & \cdot [1 + e^{(\nu-\beta)Z} - e^{-(\rho-\nu)(Z-z)}] \\ & + (\beta-\nu)\, e^{-\rho(Z-z)}\,[e^{-\nu Z} - e^{-\beta Z} - e^{-\rho(Z-z)-\nu z} + e^{-\rho(Z-z)-\beta z}] \\ & \cdot [1 - e^{-\nu(Z-z)} + e^{-\beta Z+\nu z}]\Big\}^{-1} \\ \cdot \Big\{ & \nu(e^{-\beta z} - e^{-\nu z})(e^{-\beta Z} - e^{-\nu Z})\,[e^{-\nu Z} - e^{-\beta Z} - e^{-\rho(Z-z)-\nu z} \\ & \qquad + e^{-\rho(Z-z)-\beta z}] \\ & + (\beta-\nu)\,(e^{-\beta z} - e^{-\nu z})[1 + e^{(\nu-\beta)Z} - e^{(\nu-\rho)(Z-z)} \\ & \qquad + e^{-\rho(Z-z)}(1-e^{-\nu(Z-z)}+e^{-\beta Z-\nu z})]\Big\}\end{aligned}$$

and

$$
\begin{aligned}
(86) \quad \bar{H}(z,Z) = \Big\{ & [e^{-\nu Z} - e^{-\beta Z} - e^{-\rho(Z-z)-\nu z} + e^{-\rho(Z-z)-\beta z}][1-\nu(e^{-\nu Z} - e^{-\beta Z})] \\
& + (\beta-\nu)[e^{-\nu Z} - e^{-\beta Z} - e^{-\rho(Z-z)-\nu z} + e^{-\rho(Z-z)-\beta z}] \\
& \cdot [1 + e^{(\nu-\beta)Z} - e^{-(\rho-\nu)(Z-z)}] \\
& + (\beta-\nu)\, e^{-\rho(Z-z)}[e^{-\nu Z} - e^{-\beta Z} - e^{-\rho(Z-z)-\nu z} + e^{-\rho(Z-z)-\beta z}] \\
& \cdot [1 - e^{-\nu(Z-z)} + e^{-\beta Z+\nu z}]\Big\}^{-1} \\
& \cdot \Big\{\beta(e^{-\nu z} - e^{-\beta z})(e^{-\nu Z} - e^{-\beta Z})[e^{-\rho(Z-z)-\beta Z} - e^{\rho(Z-z)-\nu Z} \\
& \qquad - e^{-\beta z} + e^{-\nu z}] \\
& + (\nu-\beta)(e^{-\nu z} - e^{-\beta z})\, [1 - e^{-\beta(Z-z)} - e^{\beta(Z-z)} + e^{-\nu Z+\beta z} \\
& \qquad + e^{\rho(Z-z)}(1 + e^{(\beta-\nu)Z})]\Big\} \quad .
\end{aligned}
$$

We then investigate the following equations

$$(87) \qquad \frac{\partial \bar{G}(z,Z)}{\partial Z} = 0 \ ,$$

$$(88) \qquad \frac{\partial \bar{G}(z,Z)}{\partial Z} = 0 \ ,$$

$$(89) \qquad \frac{\partial \bar{H}(z,Z)}{\partial z} = 0 \ ,$$

and

$$(90) \qquad \frac{\partial \bar{H}(z,Z)}{\partial z} = 0 \ ,$$

to find that they will hold if and only if

(91) $$\frac{\partial\bar{G}(z,Z)}{\partial z} = \frac{\partial\bar{H}(z,Z)}{\partial z} ,$$

and

(92) $$\frac{\partial\bar{G}(z,Z)}{\partial Z} = \frac{\partial\bar{H}(z,Z)}{\partial Z}$$

do. That is, $\bar{G}(z,Z)$ and $\bar{H}(z,Z)$ have the same local optima. With parameter situation [46] we further have that

(93) $$\lim_{z \downarrow 0} \bar{G}(z,Z) = - \lim_{z \downarrow 0} \bar{H}(z,Z) = \infty \qquad \text{for fixed } Z ,$$

(94) $$\lim_{Z \uparrow \infty} \bar{G}(z,Z) = - \lim_{Z \uparrow \infty} \bar{H}(z,Z) = \infty \qquad \text{for fixed } z ,$$

and

(95) $$\lim_{z \to Z} \bar{G}(z,Z) = - \lim_{z \to Z} \bar{H}(z,Z) = -\infty ;$$

while with [47] the results are that

(96) $$\lim_{z \downarrow 0} \bar{G}(z,Z) = - \lim_{z \downarrow 0} \bar{H}(z,Z) = -\infty \qquad \text{for fixed } Z ,$$

(97) $$\lim_{Z \uparrow \infty} \bar{G}(z,Z) = - \lim_{Z \uparrow \infty} \bar{H}(z,Z) = -\infty \qquad \text{for fixed } z ,$$

and

(98) $$\lim_{z \to Z} \bar{G}(z,Z) = - \lim_{z \to Z} \bar{H}(z,Z) = \infty .$$

Therefore there exist z^* and Z^*, where $0 < z^* < Z^* < \infty$, satisfying (91) and (92). Under [46] the uniqueness of z^* and Z^* as solution to (91) and (92) follows since $\bar{G}(z,Z)$ decreases and $\bar{H}(z,Z)$ increases in z on $(0,Z)$ for fixed Z, while $\bar{G}(z,Z)$ increases and $\bar{H}(z,Z)$ decreases in Z on (z,∞) for fixed z. Similarly, if [47] holds then $\bar{G}(z,Z)$ increases and $\bar{H}(z,Z)$ decreases in z on $(0,Z)$ for fixed Z, while $\bar{G}(z,Z)$ decreases and $\bar{H}(z,Z)$ increases in Z on (z,∞) for fixed z.

Letting our two critical numbers be the unique positive solutions z^* and Z^* $(z^* < Z^*)$ to (91) and (92), we conclude with verification of the necessary and sufficient optimality conditions. These can be summarized by the requirements

$$(99) \qquad |V_{f_*}(x,1) - V_{f_*}(x,z)| \le K \qquad \text{for all } x \in [z^*,Z^*]$$

$$(100) \qquad D_2 V_{f_*}(x,1) - \alpha V_{f_*}(x,1) + x - \alpha K \ge 0 \quad \text{for all } x \in [0,z^*] ,$$

and

$$(101) \qquad D_1 V_{f_*}(x,2) - \alpha V_{f_*}(x,2) + x - \alpha K \ge 0 \quad \text{for all } x \in [Z^*,\infty) .$$

Upon substitution of (91) and (92) into (80) and (81) with parameter restrictions [46] or [47], we find the following to be true

$$(102) \qquad V_{f_*}(x,1) - V_{f_*}(x,2) \qquad \text{increasing in } x \text{ on } [z^*,Z^*] ,$$

(103) $$V_{f_*}(z^*,1) - V_{f_*}(z^*,1) = V_{f_*}(Z^*,z) - V_{f_*}(Z^*,1) = -K ,$$

(104) $$\frac{1}{2}(\sigma_2^2-\sigma_1^2)\, V''_{f_*}(x,1) + (\mu_2-\mu_1)\, V'_{f_*}(x,1) \quad \text{increasing in } x \text{ on } [0,z^*] ,$$

(105) $$\frac{1}{2}(\sigma_2^2-\sigma_1^2)\, V''_{f_*}(z^*,1) + (\mu_2-\mu_1)\, V'_{f_*}(z^*,1) = \alpha K ,$$

(106) $$\frac{1}{2}(\sigma_1^2-\sigma_2^2)\, V''_{f_*}(x,2) + (\mu_1-\mu_2)\, V'_{f_*}(x,2) \quad \text{increasing in } x \text{ on } [Z^*,\infty) ,$$

and

(107) $$\frac{1}{2}(\sigma_1^2-\sigma_2^2)\, V''_{f_*}(Z^*,z) + (\mu_1-\mu_2)\, V'_{f_*}(Z^*,z) = \alpha K .$$

Statements (102) through (107) imply (99) through (101), and thus with parameter combinations [46] and [47] a two critical numbers policy is optimal where the critical numbers are the unique positive pair to simultaneously solve two transcendental equations.

BIBLIOGRAPHY

Arnold, L. Stochastic Differential Equations, Wiley and Sons; New York, 1974.

Ash, R.B. and Gardner, M.F. Topics in Stochastic Processes, Academic Press; New York, 1975.

Bather, J. "Diffusion Models in Stochastic Control Theory," J.R. Statist. Soc., A132; 1969, pp. 335-345.

Bellman, R. Dynamic Programming, Princeton University Press; Princeton, New Jersey, 1957.

Bensoussan, A. and Lions, J.L. "Nouvelle Methodes en Contrôle Impulsionnel," Applied Mathematics and Optimization Quarterly, Vol. 1, No. 4, 1975, pp. 289-312.

Breiman, L. Probability, Addison-Wesley; Reading, Mass., 1968.

Chernoff, H. and Petkau, A.J. "Optimal Control of a Brownian Motion," SIAM J. Appl. Math., Vol. 34, No. 4; 1978, pp. 717-731.

Constantinides, G.M. "Stochastic Cash Management with Fixed and Proportional Transaction Costs," Mgt. Sci., Vol. 22; 1976, pp. 1320-1331.

Dayanada, P. "Optimal Reinsurance," J.A.P., Vol. 7; 1970, pp. 134-156.

Dynkin, E. . Markov Processes, Vols. I and II, Academic Press, Springer-Verlag; New York, 1965.

Faddy, M.J. "Optimal Control of Finite Dams," Adv. Appl. Prob., Vol. 6; 1974, pp. 689-710.

Feller, W. "Diffusion Processes in One Dimension," Trans. Am. Math. Soc., Vol. 77; 1954, pp. 1-31.

Feller, W. "Generalized Second Order Differential Operators and Their Lateral Conditions," Ill. J. Math., Vol. 1; 1957, pp. 459-504.

Gihman, I.I. and Skorohod, A.V. Stochastic Differential Equations, Springer-Verlag, New York, 1972.

Harrison, J.M. and Taylor, A. "Optimal Control of a Brownian Storage System," Stochas. Proc. and Their Appl., Vol. 6; 1977.

Iglehart, D.L., "Diffusion Approximations in Collective Risk Theory," J.A.P., Vol. 6: 1969, pp. 285-292.

Itô, K. and McKean, H.P., Jr. Diffusion Processes and Their Sample Paths, Academic Press; New York, 1965.

Kabanov, Yu.M. "Representation of Functionals of Weiner and Poisson Processes in the Form of Stochastic Integrals," Th. Prob. Appl., Vol. 18; 1973, pp. 362-365.

Karlin, S. and Taylor, H.M. A First Course in Stochastic Processes, Academic Press; New York, 1975.

Kimura, M. "Diffusion Models in Population Genetics," J.A.P., Vol. 1; 1964, pp. 177-232.

Kunita, H. and Watanabe, S. "On Square Integrable Martingales," Nagoya Math. Journal, Vol. 31; 1967, pp. 209-245.

MacQueen, J. "A Problem in Survival," Annals of Math. Stat., Vol. 32; 1961, pp. 605-609.

MacQueen, J. "A Problem in Making Resources Last," Management Science, Vol. 11, No. 2; 1964, pp. 341-347.

Mandl, P. Analytical Treatment of One-Dimensional Markov Processes, Springer-Verlag; New York, 1968.

McKean, H.P., Jr. Stochastic Integrals, Academic Press; New York, 1969.

Mitchell, W. "Optimal Service-Rate Selection in a M/G/1 Queue," SIAM J. Appl. Math., Vol. 24; 1973, pp. 19-35.

Moran, P.A.P., "A Theory of Dams with Continuous Input and a General Release Rule," J.A.P., Vol. 6; 1969, pp. 88-98.

Pliska, S.R., "Single and Multi-Person Controlled Diffusions," Technical Report No. 13, Department of Operations Research, Stanford University, May 20, 1972.

Pliska, S.R., "A Diffusion Model for the Optimal Operation of a Reservoir System," J.A.P., Vol. 12; 1975, pp. 859-863.

Puterman, M. "A Diffusion Process Model for a Storage System," TIMS Studies in Mgt. Sci., Vol. 1; 1975, pp. 143-159.

Rath. J.H. "Controlled Queues in Heavy Traffic," Adv. Appl. Prob., Vol. 7; 1975, pp. 656-671.

Richard, S.F. "Optimal Impulse Control of a Diffusion Process with Both Fixed and Proportional Costs of Control," SIAM J. Contr. and Opt., Vol. 15; 1974, pp. 79-91.

Rudin, W. <u>Principles of Mathematical Analysis</u>, McGraw-Hill, 2nd Edition; New York, 1964.

Rudin, W. <u>Real and Complex Analysis</u>, McGraw-Hill, 2nd Edition, New York, 1974.

Zuckerman, D. "Two Stage Output Procedure of a Finite Dam," <u>J.A.P.</u>, Vol. 14,

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER #84 ✓ 14 TR-84	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) 6 SOME PROBLEMS IN THE OPTIMAL CONTROL OF DIFFUSIONS.		5. TYPE OF REPORT & PERIOD COVERED 9 Technical Report.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) 10 Diane Sheng		8. CONTRACT OR GRANT NUMBER(s) 15 N00014-76-C-0418
9. PERFORMING ORGANIZATION NAME AND ADDRESS Dept. of Operations Research Stanford University, Stanford, Calif.		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (NR-047-061)
11. CONTROLLING OFFICE NAME AND ADDRESS OPERATIONS RESEARCH PROGRAM OFFICE OF NAVAL RESEARCH Code 434 ARLINGTON, VA. 22217		11 12. REPORT DATE December 1978
		13. NUMBER OF PAGES 128 12 134 p.
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

This research was supported in part by National Science Foundation Grant ENG 75-14847 Department of Operations Research, Stanford University and issued as Technical Report No. 50.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Diffusions, Diffusion processes, Control, Optimal Control, Optimal Control of Diffusions, Stochastic Processes, Queueing Theory, Control of Queueing Systems, Optimal Control of Queueing Systems, Input-Output Systems.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Please see reverse side

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

402 766

ABSTRACT: SOME PROBLEMS IN THE OPTIMAL CONTROL OF DIFFUSIONS.

We consider a class of problems in the optimal control of one-dimensional diffusion processes, with the objective to minimize expected discounted cost over an infinite planning horizon. There are available a finite number of control modes (actions), and the state of the system changes locally like a Brownian Motion whose drift and variance depend upon the control mode being employed (but not upon the current state). There is a holding cost which is proportional to the state of the system and is independent of the control mode. In addition to these continuous costs, there are lump costs associated with a change in action. The state space may be either a finite or semi-infinite interval, and different types of boundary behavior are considered. Absorbing barriers arise in applications to collective risk and insurance, while reflecting barriers are natural for problems in the optimal control of queueing and storage systems.

When there are only two control modes, one expects an optimal policy characterized by a pair of critical numbers. For various special cases, it is shown that such an optimal policy exists, and (complicated) formulas for the critical numbers are derived.

84/50